الدكتور صواليلي صدر الدين

تحليل المعطيات

Analyse des données

دار

إن الطريقة البيداغوجية التي اتبعناه في العرض متمثلة في عرض التقنية متبوعة بمثال تطبيقي، ومجموعة من التمارين المحلولة للتوضيح، بالإضافة إلى ذلك قمنا بعرض طريقة تطبيق طريقة تحليل المعطيات على SPSS، ومن التقنيات التي عرضناه متمثلة في كل من تقنية التحليل المعطيات لجدول ذات أفراد ومتغيرات كمية، وجدول ذات أفراد ومتغيرات كيفية، وإلى كيفية التصنيف المجموعات.

د. صدر الدين صواليلي

الجزائر 2011/09/01

فهرس المحتويات

مقدمة

نسمي إحصاء الوصفي كل تحليل إحصائي يتم على جداول إحصائية ذات بعد واحد أي ذات متغيرة واحدة، أما في حالة جداول إحصائية ذات متغيرات متعددة فان التقنية أو الطريقة الملائمة لتحليلها متمثلة في التحليل الإحصائي المتعدد الأبعاد ؛ والذي هدفه استخراج من الكم الهائل من المعطيات نتائج يمكن الاعتماد عليها، وذلك بدون الاستناد على التوزيعات الإحصائية ؛ حيث أن تحليل المعطيات يمكن أن يكون عن طريق طرق كمية أخرى مثل الاقتصاد القياسي (الانحدار المتعدد والطرق التنبؤية) والتي تعتمد على التوزيعات الإحصائية أي على فرضيات متعلق بالمتغيرات، ولكن عادة عندما نتكلم عن تحليل المعطيات فإننا نقصد تلك الطرق الجبرية والهندسية الخالية من الفرضيات المتعلقة بتوزيع المتغيرات.

سمح التطور الذي عرفته برامج الإعلام من تطبيق تقنية تحليل المعطيات بسرعة وبدون أي جهد، ومن بين البرامج المستعملة لدينا برنامج: SPSS، Statistica، SAS ؛ والتي تستعمل تحت Windows وكل من برنامج (XLSTAT، StatBox) والتي هي مكملة لبرنامج Excel.

سمح الكم الهائل لبرامج الإحصائية التي ظهرت من تبسيط تطبيق طرق تحليل المعطيات، لذلك ارتأينا عرض كيفية تطبيق تحليل المعطيات على برنامج SPSS، ومن أجل فهم النظري لتحليل المعطيات، قمنا بتقسيم هذا الكتاب إلى ثلاثة فصول كما يلي :

- **الفصل الأول** : طريقة التحليل بالمركبات الرئيسية « ACP » (Analyse en Composantes Principales) تعرضنا في هذا الفصل على تقنية تحليل المعطيات الكمية، أي في حالة تجميع معطيات في جداول تحتوي على متغيرات كمية في أعمدتها وأفراد في أسطره.

- **الفصل الثاني** : طريقة التحليل بالمعاملات للتوفيقات AFC (Analyse factorielle des correspondances)، تطبق هذه التقنية على معطيات كيفية مثل الاستبيانات؛ أي أنها تطبق على جداول تحتوي على متغيرات كيفية في أعمدتها، وأفراد في أسطرها.

- **الفصل الثالث** : التصنيف التسلسلي CH (Classification Hiérarchique)، في هذا الفصل قمنا بعرض كيفية تفكيك مجموعة من الأفراد أو المتغيرات إلى مجموعات منفصلة، حيث هذه التقنية عادة ما تكون مكملة لطريقة التحليل المعطيات.

الفصل الأول

تحليل المعطيات بالمركبات الرئيسية

I- تحليل المعطيات بالمركبات الرئيسيـــة

ACP (Analyse en composante principale)

1. تمهيـــد

استعلمت هذه الطريقة لأول مرة من طرف " Karl Pearson " في 1901 ، وأول من ضمها إلى الإحصاء الرياضي " Harold Hotelling " في 1933 ، إلا أن تطبيق هذه الطريقة يعود الى استعمال وسائل الحساب الحالية (الحاسوب) ؛ تتمثل طريقة ACP في تحليل وصفي للمعطيات وبالتالي لا تستند على الاحتمالات، و لا على فرضيات بل هي طريقة هندسية تستعمل على جداول تحتوي في أعمدتها متغيرات ذات قيم كمية ، وفي أسطرها يوجد أفراد.

2. عرض وتمثيل المعطيـات

بعد جمع المعطيات المتعلقة بظاهرة ما نقوم بوضعها و تلخيصها في جداول تسمح لنا من تحليلها ، حيث عادة ما نضع المتغيرات في الأعمدة والأفراد المدروسة في الأسطر.

1.2. جدول الــمعطيات

يتم عرض ظاهرة تحتوي على أفراد معبرة عنها عن طريق متغيرات كمية في جدول مستطيل يحتوي على n سطر وp عمود كما يلي :

جدول 1 : عرض المعطيات في جدول ذات متغيرات كمية و أفراد

$$\Re^p$$

المتغيرات 1 2 ...j... p

$$\text{الأفراد} \quad \begin{matrix} 1 \\ 2 \\ \vdots \\ i \\ \vdots \\ n \end{matrix} \quad X_{n \times p} = \begin{bmatrix} x_{11} & x_{12} & \cdots x_{1j} \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots x_{2j} \cdots & x_{2p} \\ \vdots & \cdots & \cdots x_{ij} \cdots & \vdots \\ x_{n1} & x_{n2} & \cdots x_{nj} \cdots & x_{np} \end{bmatrix} \quad \Re^n$$

حيث : x_{ij} هي قيمة المتغيرة j لدى الفرد i

مثال 1: ليكن الجدول التالـي :

المتغيرات الأفراد/	X1	X2	X3	X4
A	1	2	4	2
B	2	3	6	4
C	5	2	8	5
D	3	1	10	7
F	3	2	7	5
G	4	5	7	4

نلاحظ من خلال هذا المثال البسيط أنه رغم حجمه البسيط أربع أعمدة وستة أسطر (أي خمس متغيرات كمية وستة أفراد)، فانه من الصعب علينا شرح العلاقات الموجودة بين المتغيرات والأفراد، لذا

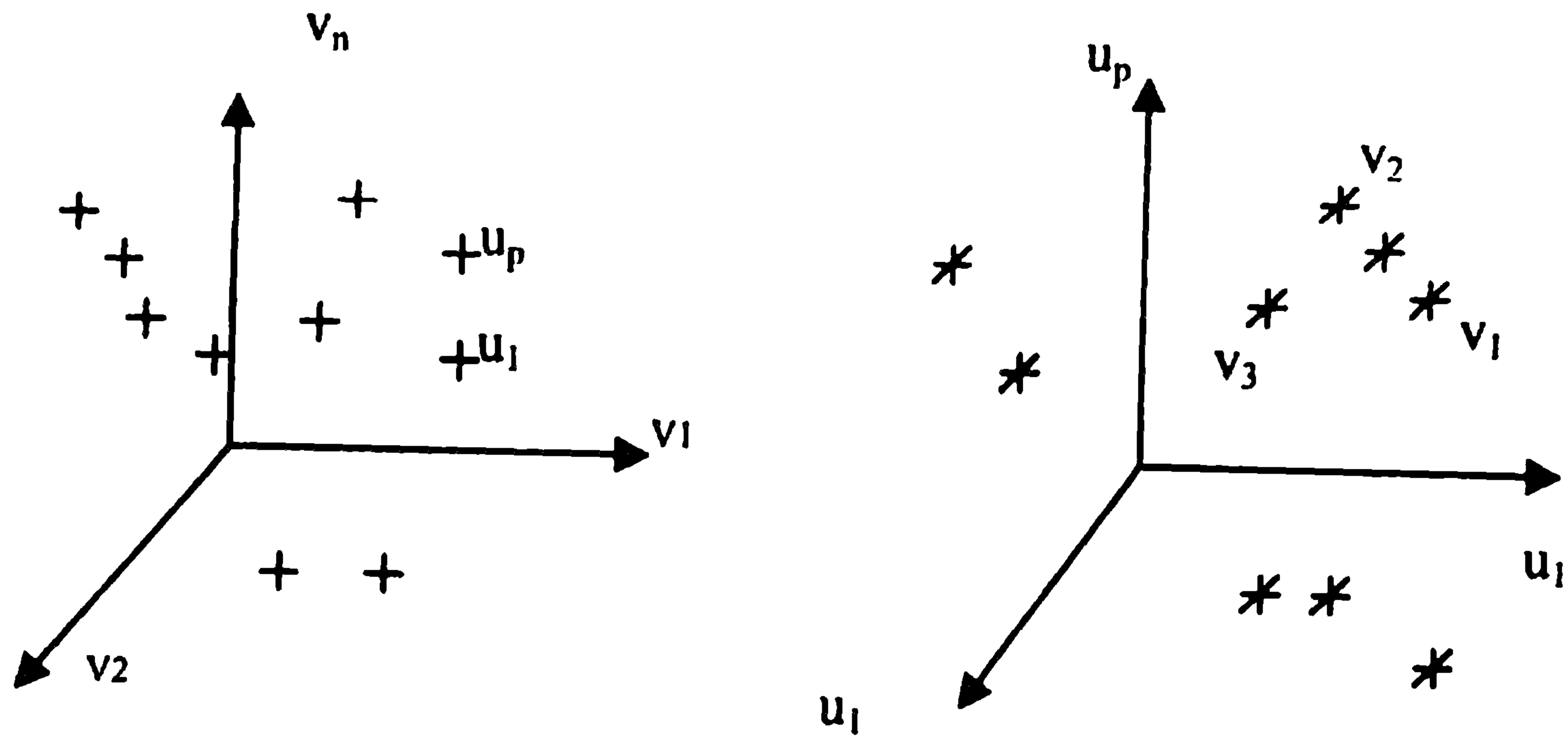
يتطلب علينا تطبيق طرق تسمح لنا من تلخيص المعلومات الموجودة في الجدول.

2.2. تمثيل المعطيات

حتى نتمكن من الحصول على الصورة لكل وحدات الجدول 1 نستعين بفضاء يأخذ كمبدأ شعاع معين من R^p على سبيل المثال الشعاع ذات المركبات المعدومة، وبالتالي كل وحدة تكون ممثل بنقطة في هذا الفضاء ؛ تسمى مجمل النقاط التي تمثل هذه الوحدات بسحابة الأفراد، وبالقيام بنفس الشيء في الفضاء R^n ؛ نسمي مجمل النقاط التي تمثل هذه الوحدات بسحابة المتغيرات.

مادام أن هذا الفضاء يمكن أن يفوق بعد من الدرجة 2 وحتى 3، فإن هدف طريقة المركبات الرئيسية هو إيجاد نظام للمحاور والمستويات، حيث إسقاط هذه السحابات من النقاط على هذه المحاور والمستويات يسمح لنا من إعادة بناء وضعية كل نقطة بالنسبة لأخرى، وبالتالي يسمح لنا من الحصول على صورة ذات أقل تشوه ؛ مع تقليص الفضاء البدائي إلى فضاء ذات بعد ذات حجم 3، كما هو موضح في الشكل 1، إلا أنه عادة ما نستعمل عند التفسير مستوي أي فضاء من درجة 2.

الشكل 1 : إسقاط الأفراد في فضاء



تتمثل طريقة المركبات الرئيسية ACP في المرور إلى جدول آخر Y ذات نفس الأفراد ، ولكن مع متغيرات جديدة أقل عددا (مادام المتغيرات البدائية مرتبطة) ؛ مع فقدان أقل ما يمكن من المعلومات :

الجدول 2 : تقليص عدد المتغيرات المفسرة للظاهرة

$$\Re^q$$

$$\begin{array}{cccc} 1 & 2 & \ldots j \ldots & q \end{array}$$

$$\begin{matrix} 1 \\ 2 \\ \vdots \\ i \\ \vdots \\ n \end{matrix} \quad \text{الأفراد} \quad Y_{n \times q} = \begin{bmatrix} x_{11} & x_{12} & \ldots x_{1j} \ldots & x_{1q} \\ x_{21} & x_{22} & \ldots x_{2j} \ldots & x_{2q} \\ \vdots & \ldots & \ldots x_{ij} \ldots & \vdots \\ x_{n1} & x_{n2} & \ldots x_{nj} \ldots & x_{nq} \end{bmatrix} \quad \Re^n$$

حيث يمكن أن نربط كل فرد i ، بالشعاع xi الفضاء الشعاعي $\Re^p$ ذات p مركبة

$$u_i = \begin{pmatrix} x_{i1} \\ x_{i2} \\ \vdots \\ x_{ip} \end{pmatrix}$$

وبإعطاء لكل فرد ثقل p_i ، حيث $p_i \geq 0$ و $\sum p_i = 1$.

نحصل على سحابة مرجحة لـ n عنصر في الفضاء $\Re^p$ (الأفراد)، وتعرف مصفوفة الثقل للأفراد بالمصفوفة القطرية التالية :

$$D = \begin{bmatrix} p_1 & & 0 \\ & \ddots & \\ 0 & & p_n \end{bmatrix}$$

نعطي لكل فرد في طريقة المركبات الرئيسية نفس الثقل أي : $p_i = \frac{1}{n}$ ، غير أنه بإمكان إعطاء للأفراد أثقال مختلفة وذلك حسب الظاهرة، ولكن يجب أن يكون مجموع الأثقال يساوي 1 ، أي $\sum_{i=1}^{n} p_i = 1$ ، وتكتب عن طريق المصفوفات بـ : $D = \frac{1}{n} I$ ، حيث I هي المصفوفة الأحادية.

مثال 2 : بأخذ المثال 1 ، فإنه يحتوي على ستة أفراد وبإعطاء نفس الثقل لكل فرد ، فإن مصفوفة الثقل هي :

$$\underline{\hat{u}}_i = \underline{a} + \sum_{k=1}^{q} y_i^k a_k$$

يكتب الإسقاط الكلي على الفضاء w بالتباين الكلي كمايلي :

$$I_w = \sum_{k=1}^{n} P_i d^2(u_i, u_{i'}) \ldots\ldots\ldots\ldots\ldots\ldots\ldots I$$

يتمثل الهدف في تصغير الكمية I_w ، وذلك باختيار المركز (المبدأ) والمحاور.

1.3. اختيار المبدأ

مع العلم أن المبدأ الصفر لا يؤدي إلى حصول على تمثيل جيد لتشتت الساحبة، لذلك نقوم بأخذ مركز ثقل السحابة كمبدأ، وعليه فإن مركز ثقل السحابة هو متوسط المتغيرات :

$$O = \begin{pmatrix} \frac{1}{n}\sum_{i=1}^{n} x_{i1} \\ \vdots \\ \frac{1}{n}\sum_{i=1}^{n} x_{il} \\ \vdots \\ \frac{1}{n}\sum_{i=1}^{n} x_{ip} \end{pmatrix}$$

يكتب عن طريق المصفوفات بـ : $o = X'D1_n$

حيث 1_n تمثل شعاع في $\mathfrak{R}^n$ ذات عناصر مساوية إلى 1.

$$D_{6X6} = \begin{bmatrix} \frac{1}{6} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{6} \end{bmatrix}$$

أي مصفوفة قطرية تحتوي على قيمة تساوي 1/6 في القطر الرئيسي، وأصفار في باقي المصفوفة. إذا من خلال ما سبق نستخلص أن طريقة المركبات الرئيسية متمثلة في البحث عن متغيرات جديدة لإسقاط الأفراد في مستوي، ومن أجل الحصول على الإسقاط يجب تحديد مبدأ (أي مركز ثقل) و مقياس لقياس المسافة مابين النقاط.

3. إسقاط جدول على مستوي

يجب اختيار وحدة القياس للنقطتين في الفضاء، من أجل إسقاط الجدول، وعليه نستعمل وحدة القياس الاقليدية التالية :

$$d^2(u_i, u_{i'}) = \sum_{j=1}^{p} (x_{ij} - x_{i'j})^2$$

حيث: ui تمثل الفرد i ، و$x$ij تمثل مركبات الفرد في الفضاء.

وحتى نتمكن من اختيار الفضاء w الذي يتم فيه الإسقاط يجب اختيار محاور من R^p المعرف بالنقطة a ؛ بحيث يكون تشويه إسقاط النقاط على هذا الفضاء أقل ما يمكن، بحيث المتغيرات الجديدة y^k عددها أقل من المتغيرة الأصلية، والتي تصبح q مركبة إسقاط للسحابة، وكل نقطة i تم إسقاطها على الفضاء تكتب كالآتي :

وبأخذ o كمركز ثقل، يتم أخذ بعين الاعتبار في هذه الحالة الجدول المركز من أجل البحث عن المحاور :

$$X_c=\begin{bmatrix} x_{11}-\frac{1}{n}\sum_{i=1}^{n}x_{i1} & x_{12}-\frac{1}{n}\sum_{i=1}^{n}x_{i2} & \cdots x_{1j}-\frac{1}{n}\sum_{i=1}^{n}x_{ij}\cdots & x_{1p}-\frac{1}{n}\sum_{i=1}^{n}x_{ip} \\ x_{21}-\frac{1}{n}\sum_{i=1}^{n}x_{i1} & x_{22}-\frac{1}{n}\sum_{i=1}^{n}x_{i2} & \cdots x_{2j}-\frac{1}{n}\sum_{i=1}^{n}x_{ij}\cdots & x_{2p}-\frac{1}{n}\sum_{i=1}^{n}x_{ip} \\ \vdots & & \cdots x_{ij}-\frac{1}{n}\sum_{i=1}^{n}x_{ij}\cdots & \vdots \\ x_{n1}-\frac{1}{n}\sum_{i=1}^{n}x_{i1} & x_{n2}-\frac{1}{n}\sum_{i=1}^{n}x_{i2} & \cdots x_{nj}-\frac{1}{n}\sum_{i=1}^{n}x_{ij}\cdots & x_{np}-\frac{1}{n}\sum_{i=1}^{n}x_{ip} \end{bmatrix}$$

إذا الشعاع المركبات المركز للوحدة ui هو :

$$U_{ci}=\begin{bmatrix} x_{i1}-\frac{1}{n}\sum_{i=1}^{n}x_{i1} \\ x_{i2}-\frac{1}{n}\sum_{i=1}^{n}x_{i2} \\ \vdots \\ x_{ij}-\frac{1}{n}\sum_{i=1}^{n}x_{ij} \\ \vdots \\ x_{ip}-\frac{1}{n}\sum_{i=1}^{n}x_{ip} \end{bmatrix}$$

تكتب المصفوفة المركزة بـ : $X_c=X-1_n O'=(I-1_n1_n'D)X$

والشعاع المركبات المركز للمتغيرات ة Vj هو :

$$V_{ci}=\begin{bmatrix} x_{1j}-\frac{1}{p}\sum_{j=1}^{p}x_{1j} \\ x_{2j}-\frac{1}{p}\sum_{j=1}^{p}x_{2j} \\ \vdots \\ x_{ij}-\frac{1}{p}\sum_{j=1}^{p}x_{ij} \\ \vdots \\ x_{nj}-\frac{1}{p}\sum_{j=1}^{p}x_{nj} \end{bmatrix}$$

مثال 3: بأخذ المثال 1:

المتغيرات الأفراد/	M1	M2	M3	M4
A	1	2	4	2
B	2	3	6	4
C	5	2	8	5
D	3	1	10	7
F	3	2	7	5
G	4	5	7	4
المجموع	18	15	42	27
المتوسط	3	2.5	7	4.5
الانحراف المعياري	1,2909	1,2583	1,8257	1,5

$$O=\begin{pmatrix} \frac{1}{6}\sum_{i=1}^{6}x_{i1} \\ \frac{1}{6}\sum_{i=1}^{6}x_{i2} \\ \frac{1}{6}\sum_{i=1}^{6}x_{i3} \\ \frac{1}{6}\sum_{i=1}^{6}x_{i4} \end{pmatrix}=\begin{pmatrix} 3 \\ 3.5 \\ 7 \\ 4.5 \end{pmatrix}$$

إذا الجدول المركز هو : $xj=Xj-Oj$

المتغيرات / الأفراد	M1	M2	M3	M4
A	-2	-0,5	-3	-2,5
B	-1	0,5	-1	-0,5
C	2	-0,5	1	0,5
D	0	-1,5	3	2,5
F	0	-0,5	0	0,5
G	1	2,5	0	-0,5

2.3. التباين الكلي لسحابة الأفــراد

يمثل التباين الكلي (Inertie) للسحابة، تشتت النقاط بالنسبة لمركزها، حيث كلما كان هذا التباين كبير كلما كانت الأفراد متشتت بالنسبة لمركز ثقلها، وكلما كان التباين صغير كلما كانت الأفراد متمركز حول مركز ثقلها ، وهو عبارة عن مجموع العناصر القطرية لمصفوفة التباين و التباين المشترك، ويمكن حسابه بـ :

$$I_O = \frac{1}{n}\sum_{i=1}^{n} d^2(O,u_i) = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{p}\left(x_{ij} - \frac{1}{n}\sum_{i=1}^{n} x_{ij}\right)^2 = \frac{1}{n}\sum_{i=1}^{n} {}^tU_{ci}U_{ci}$$

يمكن كتابة العلاقة السابقة كالآتي :

$$I_O = \frac{1}{n}\sum_{i=1}^{n} d^2(O,u_i) = \sum_{j=1}^{p}\left[\frac{1}{n}\sum_{i=1}^{n}\left(x_{ij} - \frac{1}{n}\sum_{i=1}^{n} x_{ij}\right)^2\right] = \sum_{j=1}^{p} VAR(V_j)$$

حيث (VAR (Vj هي تباين المتغيرات Vj ، وعليه فان التباين الكلي هو عبارة عن أثر مصفوفة التباين والتباين المشترك للمتغيرات؛ حيث يتم حساب مصفوفة التباين و التباين المشترك بـ :

$$V = X'DX - oo' = X'_c DX_c$$

مثال 4 : بأخذ معطيات المثال 1

مصفوفة التباين و التباين المشترك للمتغيرات هي :

$$V=\frac{1}{n}X'_cX_c=\frac{1}{6}\begin{pmatrix} -2 & -1 & 2 & 0 & 0 & 1 \\ -0.5 & 0.5 & -0.5 & -1.5 & -0.5 & 2.5 \\ -3 & -1 & 1 & 3 & 0 & 0 \\ -2.5 & -0.5 & 0.5 & 2.5 & 0.5 & -0.5 \end{pmatrix}\begin{pmatrix} -2 & -0.5 & -3 & -2.5 \\ -1 & 0.5 & -1 & -0.5 \\ 2 & -0.5 & 1 & 0.5 \\ 0 & -1.5 & 3 & 2.5 \\ 0 & -0.5 & 0 & 0.5 \\ 1 & 2.5 & 0 & -0.5 \end{pmatrix}$$

باستعمال عملية ضرب المصفوفات نحصل على :

$$V = \frac{1}{n}X'_cX_c = \frac{1}{6}\begin{pmatrix} 10 & 2 & 9 & 6 \\ 2 & 9.5 & -4 & -4.5 \\ 9 & -4 & 20 & 16 \\ 6 & -4.5 & 16 & 13.5 \end{pmatrix}$$

$$I_o = \sum_{j=1}^{4} VAR(V_j) = (10/6) + (9.5/6) + (20/6) + (13.5/6) = 8.8333$$

1.2.3. التبـايـن الكلي للأفراد بالنسبة لمحور يمر عليها

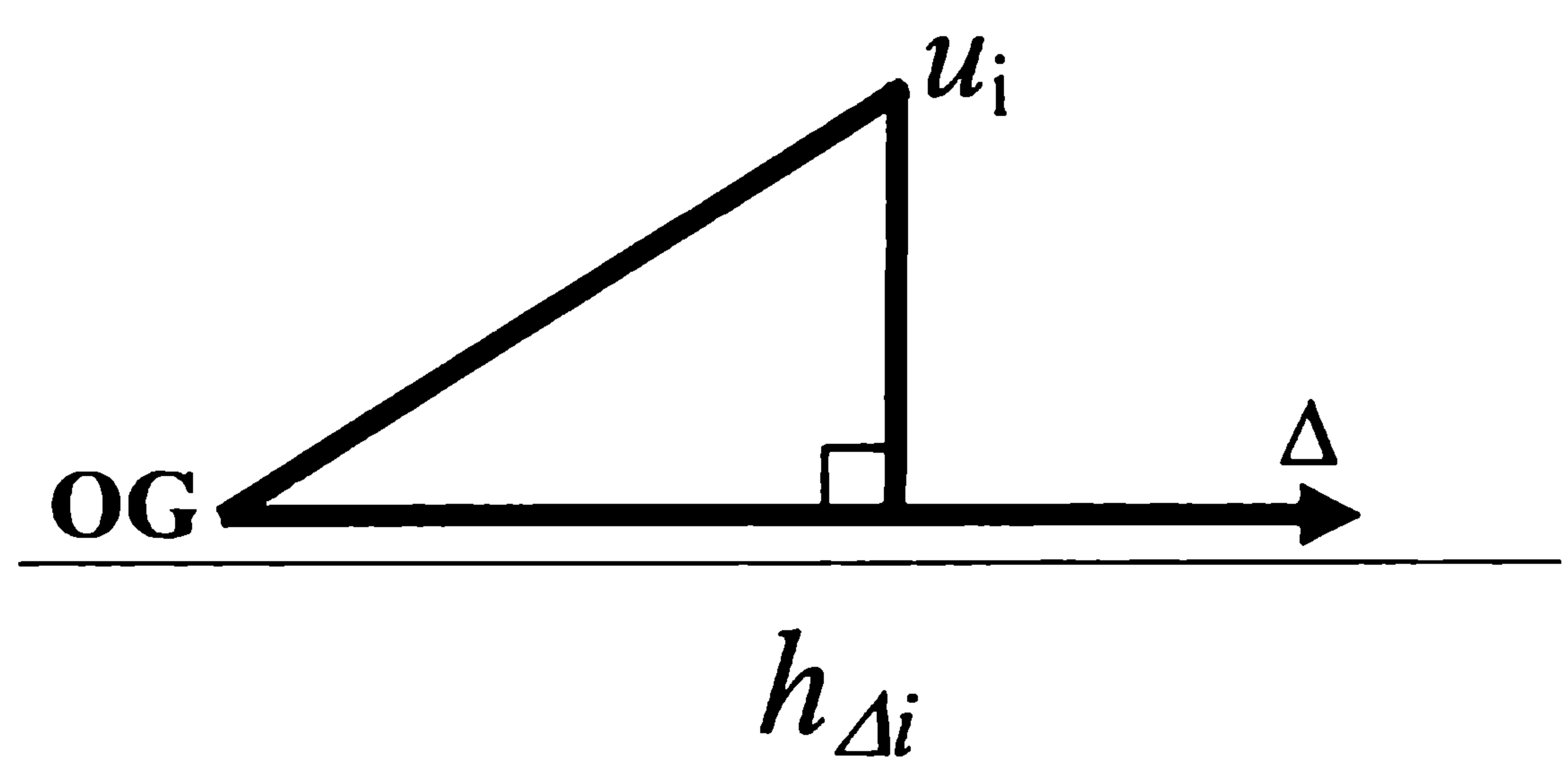
يمر تباين الأفراد بالنسبة للمحور Δ على مركز الثقل :

$$I_{\Delta} = \frac{1}{n}\sum_{i=1}^{n} d^2(h_{\Delta i}, u_i)$$

حيث $I_{\Delta i}$ هو الإسقاط العمودي لـ ui على المحور Δ.

يقيس هذا التباين اقتراب المحور من سحابة الأفراد ، أنظر الشكل 2.

الشكل 2 : اسقاط عنصر على محور



2.2.3. التبــايـن الكلــي للأفــراد بالنــسبة لفــضاء الــشعاعي الجزئــي والمركز الثقل :

ليكن الفضاء الشعاعي الجزئي w ، صيغة هذا التباين الكلي للأفراد بالنسبة لفضاء الشعاعي تعطى كالأتي :

$$I_w = \frac{1}{n}\sum_{i=1}^{n} d^2(h_{wi}, u_i)$$

حيث : h_{wi} هو الإسقاط العمودي لـ ui على الفضاء الشعاعي الجزئي w.

- تفكيك التباين الكلي: إذا قمنا بوضع W^* الفضاء المكمل العمودي للفضاء W في R^p وh_{w^*i} الاسقاط العمودي لـ ui على W* وبتطبيق علاقة فيتاغور Pytaghore ويمكن أن نكتب مايلي :

$$d^2(h_{wi},ui)+d^2(h_{w^*i},ui)=d^2(O,ui)=d^2(O,h_{wi})+d^2(O,h_{w^*i})$$

وباستعمال نظرية Huygens لدينا :

$$I_w + I_{w^*} = I_O$$

4. المحور ومركز الثقل

نقوم بالبحث عن المحور Δ1 المار من المركز O ذات التباين $I_{\Delta 1}$ أصغري والذي يقترب من سحابة الأفراد ، وذلك حتى نتمكن من عرض الأفراد في مستوي، وبالتالي فهو المحور الذي يعطي اقل تشويه للسحابة؛ بأخذ العلاقة السابقة لـ Huygens فإن تصغير $I_{\Delta 1}$ يعني تعظيم $I_{\Delta^* 1}$:

$$I_w = I_O - I_{w^*}$$

$$I_{\Delta 1} Minimum \Leftrightarrow I_{\Delta^* 1} Maximun$$

نعرف المحور Δ1 بالشعاع التوجيهي $\vec{Oa_1}$.

يتمثل النظام الذي يجب حله في :

$$Va_1 - \lambda_1 a_1 = 0 \qquad (1)$$
$$^t a_1 a_1 - 1 = 0 \qquad (2)$$

نستنتج من المعادلة (1) أن a1 هو شعاع ذاتي للمصفوفة V وبضرب طرفي المعادلة بـ $^t a_1$ نحصل على :

$$^t a_1 V a_1 - \lambda_1 {}^t a_1 a_1 = 0 \qquad (1)$$

وبتعويض المعادلة (2) في هذه المعادلة نحصل على :

$$^t a_1 V a_1 = \lambda_1$$

وعليه نستنتج أن التباين الكلي للمحور يجب أن يكون أعظمي $I_{\Delta *1}$ ، وهو يمثل بأكبر قيمة ذاتية لـ V أي λ_1 وهو محمول من طرف المحور Δ_1.

3.2.4. المحاور الثانوية

يتم البحث عن المحور الثاني Δ_2 العمودي على المحور السابق ذات التباين أصغري، باستعمال نفس الطريقة المستعملة سابقا أي تعظيم :

$$I_{\Delta_2 *} = {}^t a_2 V a_2$$

ولكن تحت القيد التالي : $^t a_2 a_1 = 0$ و $^t a_2 a_2 = 1$ ، ويعني القيد $^t a_2 a_1 = 0$ أن المحاور متعامدة.

إذا يجب ايجاد $\vec{Oa_1}$ بحيث $I_{\Delta*1}$ يكون أعظمي تحت القيد $\left\|\vec{Oa_1}\right\|^2 = 1$.

1.2.4. الكتابة الجبرية للمحور والشعاع التوجيه

$$d^2\left(O, h_{\Delta_1*i}\right) = \langle \vec{Ou_i}, \vec{Oa_1} \rangle^2 = {}^t a_1 U_{ci} {}^t U_{ci} a_1$$

باستعمال خاصية التناظر في ضرب الأشعة نحصل على :

$$I_{\Delta_1 *} = \frac{1}{n}\sum_{i=1}^{n} {}^t a_1 U_{ci} {}^t U_{ci} a_1 = {}^t a_1 \left(\frac{1}{n}\sum_{i=1}^{n} U_{ci} {}^t U_{ci} \right) a_1$$

تمثل القيمة مابين قوسين مصفوفة التباين و التباين المشترك للمتغيرات V.

$$\left\|\vec{Oa_1}\right\|^2 = {}^t a_1 a_1 \quad \text{و} \quad I_{\Delta_1 *} = {}^t a_1 V a_1$$

2.2.4. التباين الأعظمي

يتمثل في تعظيم $I_{\Delta_1 *} = {}^t a_1 V a_1$ تحت قيد ${}^t a_1 a_1 = 1$ ،

وباستعمال دالة لاقرانج :

$$o(a_1) = o(a_{11}, a_{12}, a_{13}, \ldots, a_{1p}) = {}^t a_1 V a_1 - \lambda_1 ({}^t a_1 a_1 - 1)$$

بالاشتقاق نحصل على :

$$\frac{\partial g(a_1)}{\partial a_1} = 2Va_1 - 2\lambda_1 a_1 = 0$$

وبتطبيق طريقة لاقرانج نجد أن a2 هو الشعاع الذاتي للقيمة الذاتية الثانية لمصفوفة التباين والتباين المشترك للمتغيرات، مع العلم أن هذه المصفوفة مربعة فهي تحتوي على p شعاع ذاتي تمثل أساس متعامد في R^p :

$$\begin{cases} \Delta_1 \perp \Delta_2 \perp \ldots \perp \Delta_p \\ a_1 \perp a_2 \perp \ldots \perp a_p \\ \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_p \\ I_{\Delta^*_1} \geq I_{\Delta^*_2} \geq \ldots \geq I_{\Delta^*_p} \end{cases}$$

ملاحظة : نسمي المحاور الأولى بالمحاور الأساسية.

مثال 5 : بأخذ معطيات المثال 1 لدينا مصفوفة التباين والتباين المشترك للمتغيرات هي :

$$V = \frac{1}{n} X_c' X_c = \begin{pmatrix} \frac{10}{6} & \frac{2}{6} & \frac{9}{6} & \frac{6}{6} \\ \frac{2}{6} & \frac{9.5}{6} & \frac{-4}{6} & \frac{-4.5}{6} \\ \frac{9}{6} & \frac{-4}{6} & \frac{20}{6} & \frac{16}{6} \\ \frac{6}{6} & \frac{-4.5}{6} & \frac{16}{6} & \frac{13.5}{6} \end{pmatrix}$$

من أجل حساب القيم الذاتية لهذه المصفوفة نعتمد على : $Va = \lambda a$ ، إذا من أجل الحصول على القيم الذاتية نحسب المحدد التالي :

$$|V-\lambda I_4|=\begin{vmatrix} \frac{10}{6}-\lambda & \frac{2}{6} & \frac{9}{6} & \frac{6}{6} \\ \frac{2}{6} & \frac{9.5}{6}-\lambda & \frac{-4}{6} & \frac{-4.5}{6} \\ \frac{9}{6} & \frac{-4}{6} & \frac{20}{6}-\lambda & \frac{16}{6} \\ \frac{6}{6} & \frac{-4.5}{6} & \frac{16}{6} & \frac{13.5}{6}-\lambda \end{vmatrix}=0$$

لإيجاد القيم الذاتية والأشعة الذاتية أنظر الملحق 1.

نحصل على :

$\lambda_4 = 0.04$ ، $\lambda_3 = 0.5$ $\lambda_2 = 1,94$ و $\lambda_1 = 6,36$

والأشعة الذاتية المقابل لها هي :

a1	a2	a3	a4
0,34	0,58	0,72	0,13
-0,17	0,80	-0,57	0,05
0,72	0,02	-0,24	-0,65
0,58	-0,14	0,30	0,74

يجب على طويلة كل شعاع أن تكون تساوي 1 ، اذا كان غير ذلك نقوم بقسمة كل عنصر من الشعاع الذاتي على جذر طولته حتى نحصل على الواحد.

$$\|a_1\|^2 = (0.34)^2 + (-0.17)^2 + (0.72)^2 + (0,58)^2 = 1$$

5. مساهمة المحاور في التباين الكلي

يمكن استعمال نظرية Huygens بتفكيك التباين الكلي للأفراد كمايلي :

$$I_O = I_{\Delta^*_1} + I_{\Delta^*_2} + \ldots + I_{\Delta^*_p} = \lambda_1 + \lambda_2 + \ldots + \lambda_p$$

1.5. المساهمة المطلقة

تقاس المساهمة المطلقة للمحور في التباين الكلي بـ:

$$ca(\Delta_k / I_O) = \lambda_k$$

إذا كل محور مساهمته تقاس بالقيمة الذاتية المقابل له.

2.5. المساهمة النسبية للمحور

تقاس المساهمة النسبية للمحور بالعلاقة التالية :

$$cr(\Delta_k / I_O) = \frac{\lambda_k}{\lambda_1+\lambda_2+\ldots+\lambda_p} = \frac{\lambda_k}{trace(1/nX'X)}$$

أما نسبة تمثيل المستوي تقاس بالعلاقة التالية :

$$cr(\Delta_1 \oplus \Delta_2 / I_O) = \frac{\lambda_1+\lambda_2}{\lambda_1+\lambda_2+\ldots+\lambda_p} = \frac{\lambda_1+\lambda_2}{trace(1/nX'X)}$$

- **مثال 6 : بأخذ المثال 5 :**

$$cr(\Delta_1 / I_O) = \frac{\lambda_1}{\lambda_1+\lambda_2+\lambda_3} = \frac{\lambda_1}{trace(1/nX'X)} = \frac{6.36}{8.84} = 0.7194$$

يساهم المحور الأول في تمثيل الأفراد بـ 71.94%.

أما نسبة التمثيل للمحور الثاني هي :

$$cr(\Delta_2 / I_O) = \frac{\lambda_2}{\lambda_1+\lambda_2+\ldots+\lambda_p} = \frac{1,94}{trace(\frac{X'X}{n})} = 0.2197$$

يساهم المحور الثاني في تمثيل الأفراد بـ 21.97%.

أما المستوي يمثل الأفراد بـ :

$$cr(\Delta_1 \oplus \Delta_2 / I_O) = \frac{\lambda_1+\lambda_2}{\lambda_1+\lambda_2+...+\lambda_p} = \frac{6,36+1,94}{trace(\frac{X'X}{n})} = 0,9389$$

نسبة تمثيل المستوي تقدر بـ 93,89%، أي أن فقدان المعلومات تقدر بـ 6,1 % في هذا المستوي.

ملاحظة : يتم اختيار عدد المحاور بوضع حد أدنى لعتبة التباين، وعادة ما يتم اختيارها لم تقترب من الواحد حتى يكون تمثيل الأفراد أحس ما يمكن ويمكن أخذ العتبة بالنسبة المؤية على سبيل المثال 80%.

6. تمثيل الأفراد في المحاور الجديدة (المركبات الرئيسية)

يتم تمثيل الأفراد على المستوي باسقاط سحابة النقاط u_i على المستوي الرئيسي $[b_k,b_{k'}]$، والذي يكون مشكل عادة من ثلاث مستويات رئيسية : $[b1,b2]$، $[b1,b3]$، $[b2,b3]$.

من أجل ايجاد y_{ik} مركبات الوحدة u_i على المحور Δ_k.

نقوم بإسقاط الشعاع $\vec{Ou_i}$ على هذا المحور نحصل على :

$$y_{ik} = \left\langle \vec{Ou_i}, \vec{a_k} \right\rangle = {}^t a_k U_{ci}$$

وبصفة عامة يتم حساب مركبات الأفراد في المحاور كمايلي :

$$Y_i = {}^t A U_{ci}$$

ثم نقوم بتمثيل المركبات في المستوي ذات أكبر قيم ذاتية.

مثال 7 : نأخذ معطيات المثال 6

لدينا مركبات الأفراد في المحور الأول والثاني :

a1	a2
0,34	0,58
-0,17	0,80
0,72	0,02
0,58	-0,14

إذا :

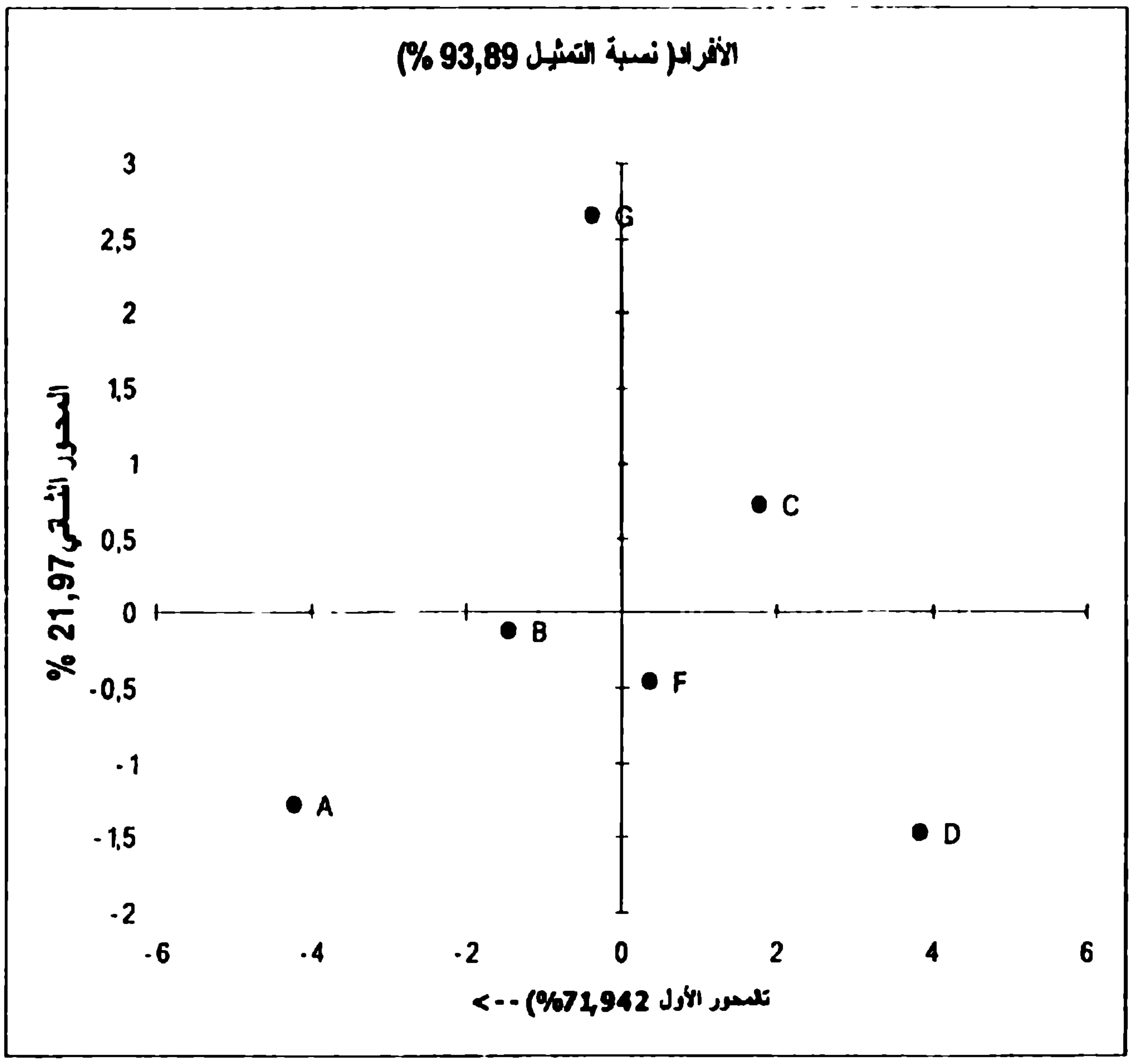
$$Y_i = {}^tA\,U_{ci} = \begin{pmatrix} 0{,}34 & -{,}17 & 0{,}72 & 0{,}58 \\ 0{,}58 & 0{,}80 & 0{,}02 & -0{,}14 \end{pmatrix} \begin{pmatrix} -2 & -1 & 2 & 0 & 0 & 1 \\ -0{,}5 & 0{,}5 & -0{,}5 & -1{,}5 & -0{,}5 & 2{,}5 \\ -3 & -1 & 1 & 3 & 0 & 0 \\ -2{,}5 & -0{,}5 & 0{,}5 & 2{,}5 & 0{,}5 & -0{,}5 \end{pmatrix}.$$

$$= \begin{pmatrix} -4{,}22 & -1{,}44 & 1{,}78 & 3{,}88 & 0{,}38 & -0{,}37 \\ -1{,}27 & -0{,}33 & 0{,}71 & -1{,}49 & -0{,}47 & 2{,}65 \end{pmatrix}$$

إذا مركبات الأفراد في المحور الأول والثاني هي :

الأفـراد	المحور الأول	المحور الثاني
A	-4,22	-1,27
B	-1,44	-0,13
C	1,78	0,71
D	3,88	-1,49
F	0,38	-0,47
G	-0,37	2,65

نسبة تمثيل الأفراد قيمتها تقترب من الصفر في المحور الثالث والرابع، إذا نقوم بتمثيل الأفراد في المحور الأول والثاني :



من أجل تمثيل الأفراد عن طريق ACP باستعمال برنامج Spss 8 أنظر الملحق 2.

1.6. النوعية النقطية للتمثيل

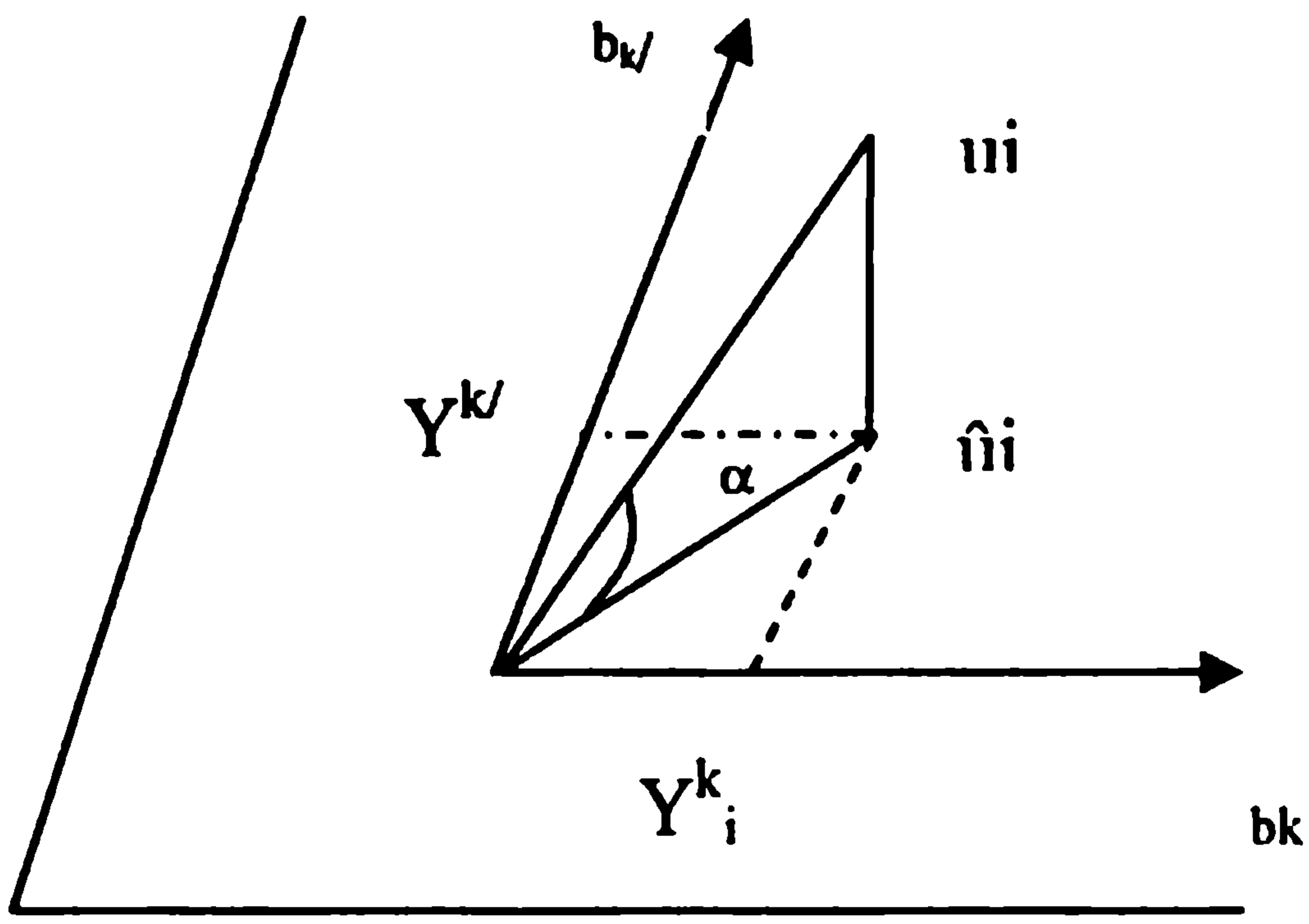
يتم قياس النوعية النقطية للتمثيل الفرد ui على المستوي $[b_k, b_{k'}]$ ، باستعمال تجب الزاوية α التي يشكلها مع المستوي، نسمي الكمية $\cos^2(\alpha_{kk'})$ ، بمساهمة النسبية للمستوي $[b_k, b_{k'}]$ في تمثيل الأفراد.

بحيث يتم حساب تجب الزاوية بالنسبة للمحور الأول باستعمال العلاقة التالية :

$$\cos^2(\alpha_{ik}) = \frac{\left\langle \vec{Ou_i}, \vec{Oa_k} \right\rangle^2}{\left\| \vec{Ou_i} \right\|^2} = \frac{{}^t a_k U_{ci} {}^t U_{ci} a_k}{{}^t U_{ci} U_{ci}} = \frac{\left[\sum_{j=1}^{p} \left(x_{ij} - \frac{1}{n} \sum_{i=1}^{n} x_{ij} \right) a_{kj} \right]^2}{\sum_{j=1}^{p} \left(x_{ij} - \frac{1}{n} \sum_{i=1}^{n} x_{ij} \right)}$$

إذا يتم حساب تجب هذه الزاوية بالنسبة للمستوي [bk,bk/] كالأتي :

$$\cos^2(\alpha_{kk'}) = \cos^2(\alpha_{ik}) + \cos^2(\alpha_{ik'})$$



مثال 8 : بأخذ نفس المثال السابق 7.

نقوم بحساب بالنسبة **للمحور الأول** نسبة مساهمة هذا المحور في تمثيل الأفراد كالأتي :

$$\cos^2(\alpha_{i1}) = \frac{\left\langle \vec{Ou_i}, \vec{Oa_1} \right\rangle^2}{\left\| \vec{Ou_i} \right\|^2} = \frac{{}^t a_1 U_{ci} {}^t U_{ci} a_1}{{}^t U_{ci} U_{ci}} = \frac{\left[\sum_{j=1}^{p} \left(x_{ij} - \frac{1}{n} \sum_{i=1}^{n} x_{ij} \right) a_{1j} \right]^2}{\sum_{j=1}^{p} \left(x_{ij} - \frac{1}{n} \sum_{i=1}^{n} x_{ij} \right)^2}$$

نحن لدينا :

al	0,34	-0,17	0,72	0,58

المتغيرات / الأفراد	M1	M2	M3	M4
A	-2	-0,5	-3	-2,5
B	-1	0,5	-1	-0,5
C	2	-0,5	1	0,5
D	0	-1,5	3	2,5
F	0	-0,5	0	0,5
G	1	2,5	0	-0,5

$$\cos^2(\alpha_{a1})=\frac{{}^t a_1 U_{ca}\,{}^t U_{ca} a_1}{{}^t U_{ca} U_{ca}}=\frac{\left[\sum_{j=1}^{p}\left(x_{aj}-\frac{1}{n}\sum_{i=1}^{n}x_{aj}\right)a_{1j}\right]^2}{\sum_{j=1}^{p}\left(x_{aj}-\frac{1}{n}\sum_{i=1}^{n}x_{aj}\right)^2}$$

$$=\frac{[(-2)(0.34)+(-0{,}17)(-0.5)+(-3)(0.72)+(-2{,}5)(0{,}58)]^2}{(-2)^2+(-0.5)^2+(-3)^2+(-2{,}5)^2}=\frac{17{,}68}{19.5}=0.9067$$

باستعمال نفس الطريقة للأفراد الأخرى نحصل على الجدول التالي :

	البسط	المقام	$Cos^2(\alpha 1)$
A	17,68	19,5	0,91
B	2,05	2,5	0,82
C	3,15	5,5	0,57
D	14,93	17,5	0,85
F	0,14	0,5	0,28
G	0,14	7,5	0,02

أي كل من الفرد A، B، D و C ممثلة جيدا في المحور الأول.

تحسب نسبة مساهمة المحور الثاني في تمثيل الأفراد كما يلي :

$$\cos^2(\alpha_{i2}) = \frac{\left\langle \vec{Ou_i}, \vec{Oa_2} \right\rangle^2}{\left\| \vec{Ou_i} \right\|^2} = \frac{{}^t a_2 U_{ci} {}^t U_{ci} a_2}{{}^t U_{ci} U_{ci}} = \frac{\left[\sum_{j=1}^{p} \left(x_{ij} - \frac{1}{n} \sum_{i=1}^{n} x_{ij} \right) a_{2j} \right]^2}{\sum_{j=1}^{p} \left(x_{ij} - \frac{1}{n} \sum_{i=1}^{n} x_{ij} \right)^2}$$

لدينا :

a2	0,58	0,80	0,02	-0,14

$$\cos^2(\alpha_{a1}) = \frac{{}^t a_1 U_{ca} {}^t U_{ca} a_1}{{}^t U_{ca} U_{ca}} = \frac{\left[\sum_{j=1}^{p} \left(x_{aj} - \frac{1}{n} \sum_{i=1}^{n} x_{aj} \right) a_{1j} \right]^2}{\sum_{j=1}^{p} \left(x_{aj} - \frac{1}{n} \sum_{i=1}^{n} x_{aj} \right)^2}$$

$$= \frac{[(-2)(0.34)+(-0,17)(-0.5)+(-3)(0.72)+(-2,5)(0,58)]^2}{(-2)^2+(-0.5)^2+(-3)^2+(-2,5)^2} = \frac{17,68}{19.5} = 0.9067$$

إذا نحصل على :

	البسط	المقام	Cos2 (α2)
A	1,61	19,5	0,08
B	0,02	2,5	0,01
C	0,50	5,5	0,09
D	2,22	17,5	0,13
F	0,22	0,5	0,44
G	7,02	7,5	0,94

من خلال هذا الحساب فان نسبة التمثيل للأفراد في هذا المستوي جيدة لكل من الفرد A و B ، D ، و G ، حيث أنها تفوق 80% ؛ وتمثيل من F و C متوسط حيث أنه يفوق 50 % ؛ أما في المحور الأول فإن كل من الفرد A ، B ، وD ممثلة جيدا ؛ أما في المحور الثاني فإن الفرد G هو الممثل جيدا تقدر نسبة إسقاطها بـ 94 %.

2.6. تفسير المركبات الرئيسية

يستلزم حساب المساهمة المطلقة للأفراد في المحاور الرئيسية ، من أجل تفسير المركبات الرئيسية.

يعطي تباين المحور بـ $I_{\Delta^* k}$ ويساوي $\frac{1}{n}\sum_{i=1}^{n} d^2(h_{\Delta ki}, O)$

وبتالي فإن المساهمة المطلقة لـ ui على هذا التباين هي :

$$ca(ui / \Delta_k) = \frac{1}{n}\sum_{i=1}^{n} d^2(h_{\Delta ki}, O)$$

$$ca(ui / \Delta_k) = \frac{\frac{1}{n}\sum_{i=1}^{n} d^2(h_{\Delta ki}, O)}{I_{\Delta^* k}} = \frac{\frac{1}{n}\left\langle \vec{Ou_i}, \vec{Oa_k} \right\rangle^2}{\lambda_k} = \frac{\frac{1}{n}{}^t a_k U_{ci} {}^t U_{ci} a_k}{\lambda_k}$$

نستنتج من خلال العلاقة السابقة أن :

$$\sum_{i=1}^{n} ca(ui / \Delta_k) = 1$$

مثال 9: نقوم بحساب المساهمة المطلقة للأفراد في بناء كل محور للمثال 8 ؛ نحن لدينا إحداثيات الأفراد في كل محور ، وعليه فانه

يتم حساب المساهمة المطلقة بترتيبع المركبات، وبقسمة كل واحد منها على الثقل المصفوفة 1/n ؛ ثم على القيمة الذاتية المقابلة لها فنحصل على :

المساهمة المطلقة في المحور الأول نقوم بالقسمة على ($\lambda 1 = 6{,}36$)	مربع المحور الأول على ثقل كل فرد في المصفوفة 1/6	مربع المحور الأول	المحور الأول	الأفراد
0,4476	2,84680	17,8084	-4 ,22	A
0,0543	0,3456	2,0736	1,44	B
0,0830	0,5280	3,1684	1,78	C
0,3944	2,5090	15,0544	3,88	D
0,0037	0,0240	0,1444	0,38	F
0,0035	0,0228	0,1369	-0,37	G
1	المساهمة الكلية للأفراد في المحور الأول			

عند جمع المساهمة المطلقة لكل الأفراد نحصل على المساهمة الكلية تساوي 1 ، أي التمثيل الأفراد هو 100% في كل محور.

المساهمة المطلقة في المحور الثاني نقوم بالقسمة على (λ2 = 1,94)	الأفراد
13,85	A
0,14	B
4,32	C
19,07	D
1,89	F
60,32	G
100%	المساهمة الكلية للأفراد في المحور الثاني

المساهمة المطلقة في المحور الأول نقوم بالقسمة على (λ1 = 6,32)	الأفراد
44,76	A
5,43	B
8,30	C
39,44	D
0,37	F
0,35	G
100%	المساهمة الكلية للأفراد في المحور الأول

نستنتج من خلال المساهمة النسبية للأفراد أن المحور الأول محدد أو مكون من طرف الأفراد A ، D بأكثر من 80%، أما المحور الثاني فهو مكون من الفرد G بـ 60,32% وكذلك من الفرد D وA بدرجة أقل 19,07% و13,85% على الترتيب.

7. تمثيل المتغيرات في المحاور الجديدة

يمكن القيام بتمثيل المتغيرات باستعمال نفس الخطوات التي قمنا بها في تمثيل الأفراد وذلك على الفضاء R^n بدلا من R^p ؛ غير أننا نستعمل العملية المتمثلة في تمثيل الأفراد في فضاء المتغيرات القديمة (الأولى)، ونقوم بتغير الأساس في هذا الفضاء، المحاور الجديدة هي عبارة عن التركيبات الخطية للمحاور القديمة، ويمكن أن نعتبرها كمتغيرات جديدة مركبة خطيا بالمتغيرات الأولى.

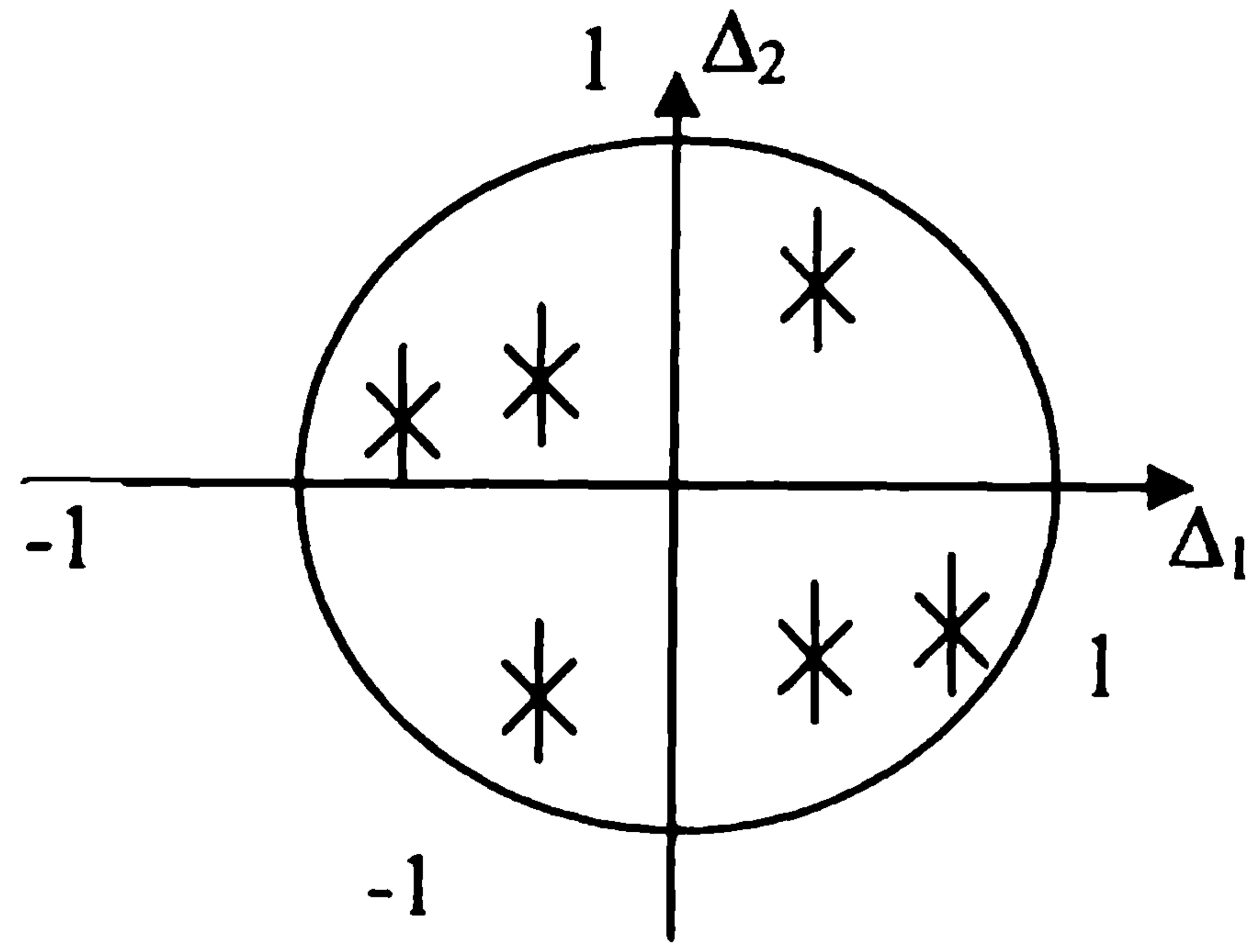
نسمي هذه المتغيرات الجديدة، بالمركبات الرئيسية، نضع :
$z_1, z_2, \dots, z_k, \dots, z_p$ المركبات الرئيسية، Z_k هي المتغيرات الجديدة المتعلقة بالمحور Δ_k :

$$z_k = \sum_{j=1}^{p} a_{kj} V_{cj} = X_c a_k$$

وبصفة عامة : $Z = [z_1, z_2, \dots, z_k, \dots, z_p] = X_c A$

الهدف هنا هو معرفة العلاقة مابين المتغيرات القديمة والمتغيرات الجديدة عن طريق الارتباط، ويتم تمثيل المتغيرات القديمة عن طريق أخذ مركبات المتغيرات القديمة (السابقة) أي معاملات ارتباطها مع المتغيرات الجديدة ؛ فنحصل على ما يسمى بـ " دائرة الارتباطات" وهذا لكون الارتباط محصور مابين -1 و1.

الشكل 3 : تمثيل المتغيرات في دائرة الارتباطات



معاملات التباين والتباين المشترك للمتغيرات الجديدة معطاة بـ :

$$Var(z_k) = \frac{1}{n} {}^t a_k {}^t Z_c Z_c a_k = {}^t a_k \Sigma a_k$$

و:

$$Cov(z_k, V_{cj}) = \frac{1}{n} {}^t a_k {}^t X_c V_{cj} = \frac{1}{n} {}^t a_k {}^t X_c X_c \begin{bmatrix} 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \end{bmatrix}$$

$$Cov(z_k, V_{cj}) = {}^t a_k \Sigma \begin{bmatrix} 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \end{bmatrix} = \lambda_k {}^t a_k \begin{bmatrix} 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \end{bmatrix} = \lambda_k a_{kj}$$

وفي الأخير لدينا الارتباط كالأتي: $R_{Z,V} = D_{\sqrt{\lambda}} {}^t A D_{\frac{1}{\sigma}}$

حيث : $Cor(z_k, V_{cj}) = \sqrt{\lambda_k} \dfrac{a_{kj}}{\sqrt{Var(V_j)}}$

مثال 10 : بأخذ نتائج المثال 9 نقوم بحساب مركبات المتغيرات لدينا :

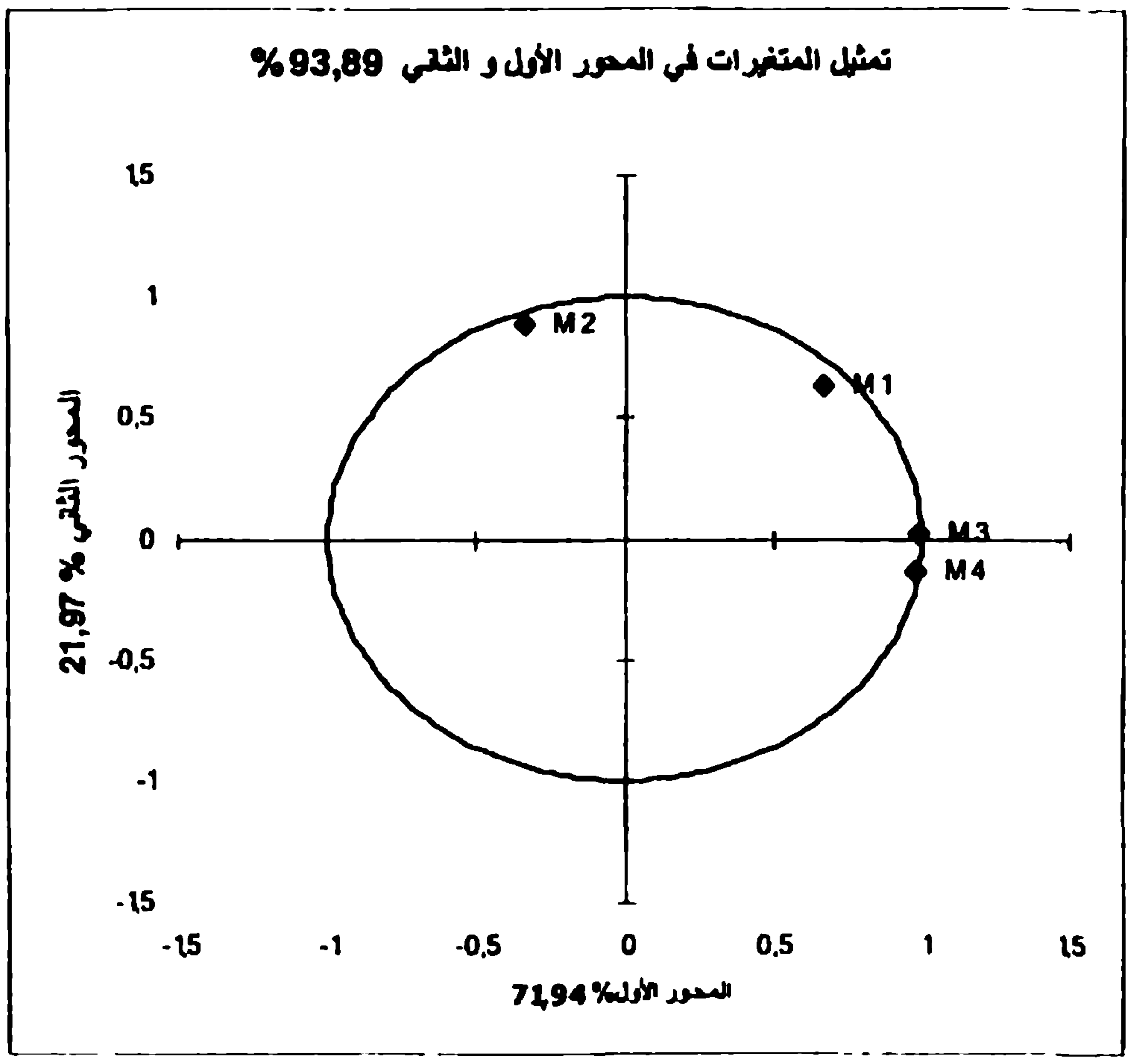
$$D_{\sqrt{\lambda}} = \begin{pmatrix} \sqrt{6,36} & 0 \\ 0 & \sqrt{1,94} \end{pmatrix} D_{\frac{1}{\sigma}} = \begin{pmatrix} \frac{1}{1,290} & & & \\ & \frac{1}{1,2583} & & \\ & & \frac{1}{1,8257} & \\ & & & \frac{1}{1,5} \end{pmatrix} A = \begin{pmatrix} 0,34 & -0,17 & 0,72 & 0,58 \\ 0,58 & 0,80 & 0,02 & -0,14 \end{pmatrix}$$

$$R_{zv}=\begin{pmatrix}\sqrt{6{,}36} & 0\\ 0 & \sqrt{1{,}94}\end{pmatrix}\begin{pmatrix}0{,}34 & -0{,}17 & 0{,}72 & 0{,}58\\ 0{,}58 & 0{,}80 & 0{,}02 & -0{,}14\end{pmatrix}\begin{pmatrix}\frac{1}{1{,}290} & & & \\ & \frac{1}{1{,}2583} & & \\ & & \frac{1}{1{,}8257} & \\ & & & \frac{1}{1{,}5}\end{pmatrix}=\begin{pmatrix}0{,}66 & -0{,}34 & 0{,}99 & 0{,}97\\ 0{,}62 & 0{,}88 & 0{,}01 & -0{,}129\end{pmatrix}$$

إذا إحداثيات المتغيرات في المحاور :

المتغيرات	المحور الأول	المحور الثاني
M1	0,66	0,62
M2	-0,34	0,88
M3	0,99	0,01
M4	0,97	-0,12

التمثيل البياني للمتغيرات في المستوي (المحور الأول والثاني):



8. تحليل المعطيات بالمركبات الرئيسية المعيارية ACP normée

إن هدف تحليل المعطيات هو تمثيل في مستوي الأفراد والمتغيرات في مستوي والقيام بالتفسير والتحليل، ولكن إذا كانت هذه المتغيرات ذات أحجام مختلفة أي ذات وحدات مختلفة، فإن الإسقاط في المستوي سيؤدي إلى تحليل مشوه للحقيقة لذلك يجب تحويل الجدول المراد تحليله.

نستعمل في حالة متغيرات ذات وحدات قياس مختلفة تحليل المعطيات المعيارية، تتمثل هذه الطريقة في إزالة آثر وحدة القياس، ويتم ذلك بقسمة كل متغيرة ممركز على انحراف معيارها في الجدول المركز، وبذلك وحدة القياس المستعملة هي :

$$d^2(u_i, u_{i'}) = \sum_{j=1}^{p} \frac{1}{\sigma_j^2}(x_{ij} - x_{i'j})^2$$

تتم عملية البحث عن المحاور الأساسية والمركبات الرئيسية من خلال هذا المقياس عن طريق مصفوفة الارتباطات بدلا من مصفوفة التباينات والتباينات المشتركة.

مثال 11 : لتكن المعطيات المتعلقة بثلاث متغيرات ذات وحدات القياس مختلفة، وللأربع أفراد كما يلي :

I	X1	X2	X3
A	50	5	1
B	43	6	3
C	62	7	2
D	30	8	3

يتم تمثيل الأفراد والمتغيرات في المستوي ومساهمة الأفراد المطلقة والنسبية، كمالي :

I	X1	X2	X3
A	50	5	1
B	43	6	3
C	62	7	2
D	30	8	3
المتوسط	46,25	6,5	2,25
الانحراف المعياري	11,5839	1,1180	0,8291

1- مع العلم أن وحدة القياس مختلفة نستعمل الانحراف المعياري كوحدة قياس.

2- الجدول المركز المعياري هو :

$$\frac{1}{\sigma_j}(x_{ij}-x_{ij})$$

أي :

X

x1	x2	x3
0,3237	-1,3416	-1,5076
-0,2806	-0,4472	0,9045
1,3596	0,4472	-0,3015
-1,4028	1,3416	0,9045

مصفوفة الارتباطات هـي :

1/4 (x/x) =

1	-0,3957	-0,6052
-0,3957	1	0,6742
-0,6052	0,6742	1

القيم الذاتية لمصفوفة الارتباطات هي :

λ1 = 2,12 λ 2 = 0,61 λ 3 = 0,27

والأشعة الذاتية المقابلة لها هـي :

a1	a2	a3
-0,54	0,76	0,36
0,56	0,64	-0,52
0,63	0,07	0,78

نسبة التمثيل في المستوي هي :

$$cr(\Delta_1 \oplus \Delta_2 / I_G) = \frac{\lambda_1 + \lambda_2}{\lambda_1 + \lambda_2 + ... + \lambda_p} = \frac{2.12+0.61}{trace(1/nX'X)} = 0.91$$

%91

1. التمثيل الأفراد في المستوي

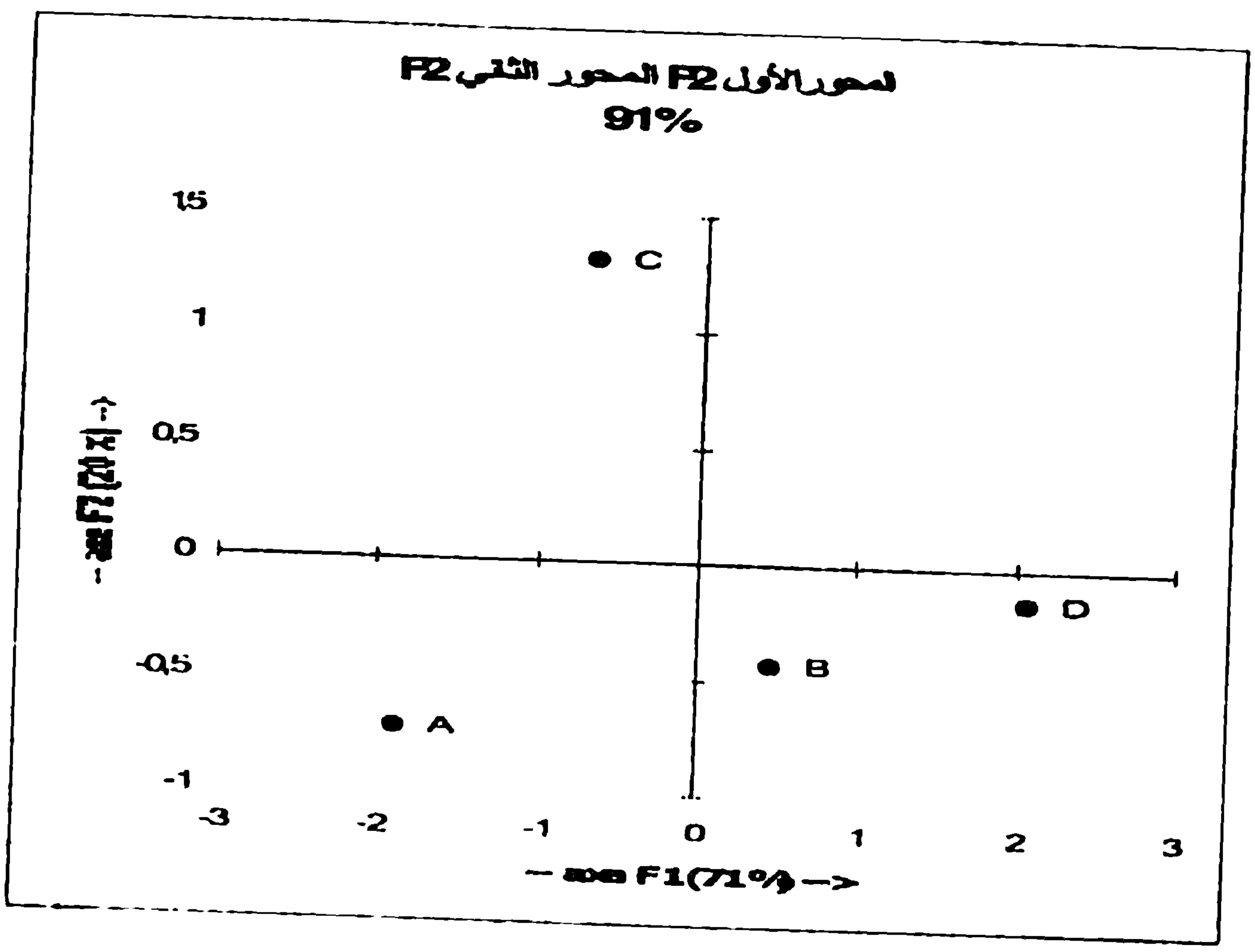
$$Y_i = {}^tAU_{ci} = \begin{pmatrix} -0.54 & 0.56 & 0.63 \\ 0.76 & 0.64 & 0.07 \\ 0.36 & -0.52 & 0.78 \end{pmatrix} \begin{pmatrix} 0.3237 & -0.2806 & 1.3596 & -1.4028 \\ -1.3416 & -0.4472 & 0.4472 & 1.3416 \\ -1.5076 & 0.9045 & -0.3015 & 0.9045 \end{pmatrix}$$

$$= \begin{pmatrix} -1.88 & 0.47 & -0.67 & 2.08 \\ -0.73 & -0.43 & 1.30 & -0.14 \\ -0.36 & -0.83 & 0.02 & -0.50 \end{pmatrix}$$

إذا مركبات الأفراد في المحاور الثالث هي :

	المحور 1	المحور 2	المحور 3
A	-1,88	-0,73	-0,36
B	0,47	-0,43	0,83
C	-0,67	1,30	0,02
D	2,08	-0,14	-0,50

إن تمثيل الأفراد في المستوي المكون من المحور الأول والثاني معطى ب:



أ. تمثيل المتغيرات : من مصفوفة الارتباطات.

$$R_{Z,V} = D_{\sqrt{\lambda}}\,{}^{t}A$$

$\lambda 1 =$ 2,12 $\lambda 2 =$ 0,61 $\lambda 3 =$ 0,27

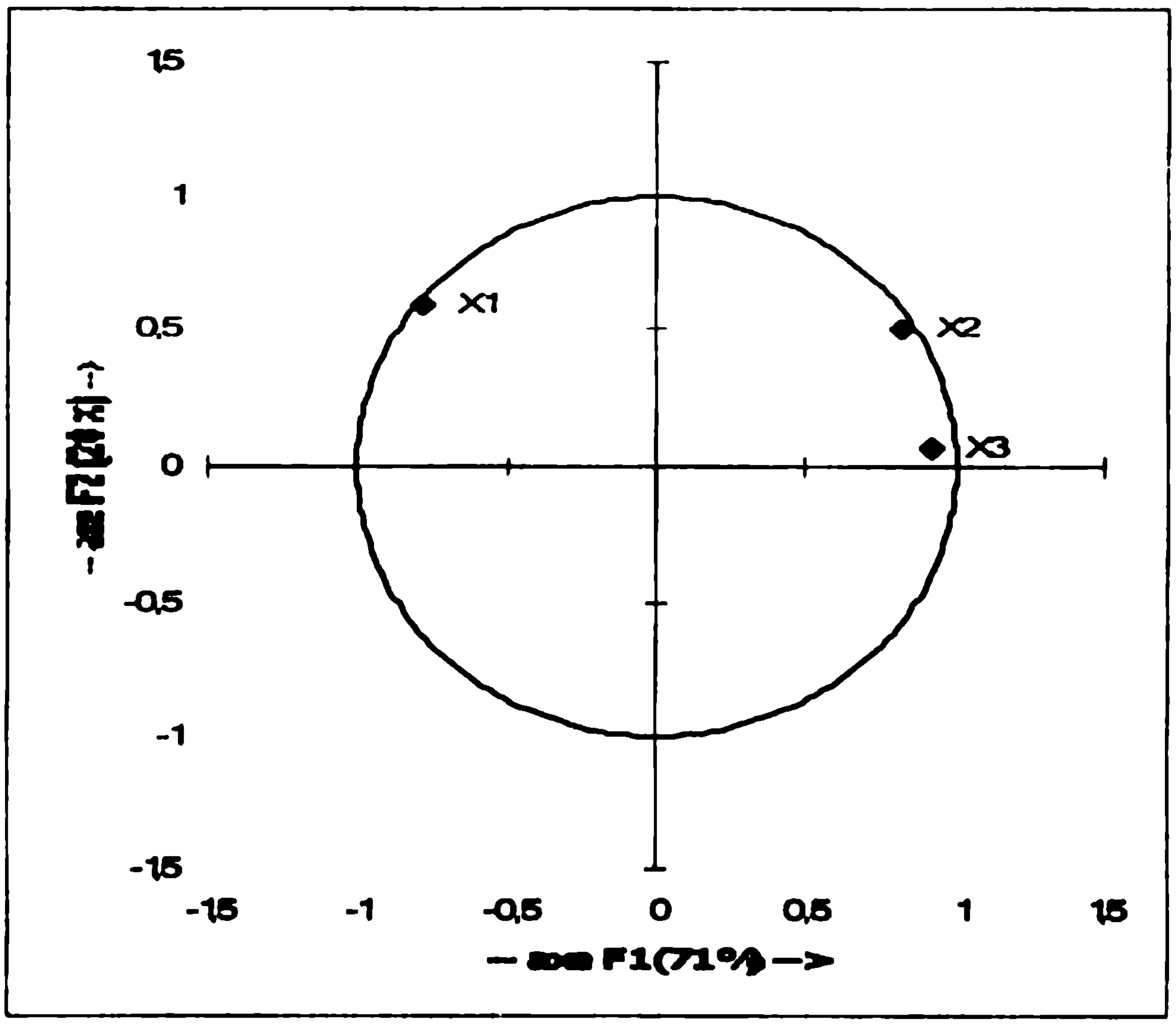
a1	a2	a3
-0,54	0,76	0,36
0,56	0,64	-0,52
0,63	0,07	0,78

$$D\sqrt{\lambda} = \begin{pmatrix} \sqrt{2.12} & 0 & 0 \\ 0 & \sqrt{0.61} & 0 \\ 0 & 0 & \sqrt{0.27} \end{pmatrix} \quad {}^{t}A = \begin{pmatrix} -0.54 & 0.56 & 0.63 \\ 0.76 & 0.64 & 0.07 \\ 0.36 & -0.52 & 0.78 \end{pmatrix}$$

$$R_{Z,V} = \begin{pmatrix} \sqrt{2.120} & 0 & 0 \\ 0 & \sqrt{0.61} & 0 \\ 0 & 0 & \sqrt{0.27} \end{pmatrix} \begin{pmatrix} -0.54 & 0.56 & 0.63 \\ 0.76 & 0.64 & 0.07 \\ 0.36 & -0.52 & 0.78 \end{pmatrix} = \begin{pmatrix} -0.79 & 0.82 & 0.92 \\ 0.59 & 0.50 & 0.05 \\ 0.19 & -0.27 & 0.41 \end{pmatrix}$$

إذا إحداثيات المتغيرات في المحاور هـي :

	المحور F1	المحور F2	المحور F3
X1	-0,79	0,59	0,19
X2	0,82	0,50	-0,27
X3	0,92	0,05	0,41



النوعية النقاطية للتمثيل :

بالنسبة للمحور الأول نقوم بحساب نسبة مساهمة هذا المحور في تمثيل الأفراد كالأتي :

$$\cos^2(\alpha_{i1})=\frac{\left\langle \vec{G}_{y},\vec{G}_{q}\right\rangle^2}{\left\|\vec{G}_{y}\right\|^2}=\frac{{}^{t}a_1 U_{ci}\,{}^{t}U_{ci} a_1}{{}^{t}U_{ci} U_{ci}}=\frac{\left[\sum_{j=1}^{p}\left(x_{ij}-\frac{1}{n}\sum_{i=1}^{n}x_{ij}\right)a_{1j}\right]^2}{\sum_{j=1}^{p}\left(x_{ij}-\frac{1}{n}\sum_{i=1}^{n}x_{ij}\right)^2}$$

نحن لدينا :

a1	-0.54	0.56	0.63

	X1	X2	X3
A	0,3237	-1,3416	-1,5076
B	-0,2806	-0,4472	0,9045
C	1,3596	0,4472	-0,3015
D	-1,4028	1,3416	0,9045

$$\cos^2(\alpha_{a1}) = \frac{{}'a_1 U_{ca} {}'U_{ca} a_1}{{}'U_{ca} U_{ca}} = \frac{\left[\sum_{j=1}^{p}\left(x_{aj} - \frac{1}{n}\sum_{i=1}^{n} x_{aj}\right) a_{1j}\right]^2}{\sum_{j=1}^{p}\left(x_{aj} - \frac{1}{n}\sum_{i=1}^{n} x_{aj}\right)^2} = \frac{[(0.3237)(-0.43)+(-1.3416)(0.56)+(-1.5076)(-0.63)]^2}{(-0.3237)^2+(-1.3416)^2+(-1.5076)^2}$$

$$= \frac{3.604}{4.1775} = 0.8628$$

	البسط	المقام	cos2(α 1)
a	3,6046	4,1775	0,8628
b	0,2164	1,0968	0,1973
c	0,4431	2,1395	0,2071
d	4,3155	4,5860	0,9410

بالنسبة للمحور الثاني نقوم بحساب نسبة مساهمة المحور الثاني في تمثيل الأفراد :

$$\cos^2(\alpha_{i2}) = \frac{\left\langle \vec{Gu_i}, \vec{Ga_2} \right\rangle^2}{\left\| \vec{Gu_i} \right\|^2} = \frac{{}'a_2 U_{ci} {}'U_{ci} a_2}{{}'U_{ci} U_{ci}} = \frac{\left[\sum_{j=1}^{p}\left(x_{ij} - \frac{1}{n}\sum_{i=1}^{n} x_{ij}\right) a_{2j}\right]^2}{\sum_{j=1}^{p}\left(x_{ij} - \frac{1}{n}\sum_{i=1}^{n} x_{ij}\right)^2}$$

لدينا :

a2	0.76	0.64	0.07

إذا نحصل على :

	البسط	المقام	cos2(α 2)
a	0,5157	4,1775	0,1235
b	0,1902	1,0969	0,1734
c	1,6859	2,1395	0,7880
d	0,0208	4,5860	0,0045

إذا المساهمة النسبية في المستوي [b_1, b_2] لكل فراد هي :

	cos2(α1)	cos2(α2)	المستوي	%
A	0,84	0,12	0,9658	96,58
B	0,20	0,17	0,3756	37,56
C	0,21	0,79	1,00	100
D	0,94	0,00	0,9467	94,67

من خلال هذا الجدول نلاحظ أن B ممثلة في المستوي بـ 37.56% أي تمثيل ضعيف، أما الأفراد الأخرى ممثلة جيدا.

المساهمة المطلقة للفرد في المحور :

$$ca(ui/\Delta_k)=\frac{\frac{1}{n}\sum_{i=1}^{n}d^2(h_{\Delta ki},G)}{I_{\Delta^* k}}=\frac{\frac{1}{n}\left\langle \vec{Gu_i},\vec{Ga_k}\right\rangle^2}{\lambda_k}=\frac{\frac{1}{n}{}^t a_k U_{ci}{}^t U_{ci} a_k}{\lambda_k}$$

تحليل المعطيات

	المحور 1	المحور 2	المحور 3
A	-1,88	2-0,7	-0,36
B	0,47	-0,43	0,83
C	-0,67	1,30	0,02
D	2,08	-0,14	-0,50

الأفراد	المحور الأول	مربع المحور الأول	مربع المحور الأول على ثقل كل فرد في المصفوفة ¼	المساهمة المطلقة في المحور الأول نقوم بالقسمة على (2,12=1λ)
A	-1,8769	3,5228	0,8807	**0,4154**
B	0,4652	0,2164	0,0541	**0,0255**
C	-0,6657	0,4432	0,1108	**0,0523**
D	2,0774	4,3156	1,0789	**0,5089**
المساهمة الكلية للأفراد في المحور الأول				**1,00**

الأفــراد	المحور الثاني	مربع المحور الثاني	مربع المحور الثاني على ثقل كل فرد في المصفوفة 4/1	المساهمة المطلقة في المحور الثاني نقوم بالقسمة على 0.61=2λ)
A	-0,7267	0,5281	0.1320	**0,2164**
B	-0,4333	0,1877	0.0469	**0,0769**
C	1,3022	1,6958	0.4240	**0,6950**
D	-0,1422	0,0202	0.0051	**0,0083**
المساهمة الكلية للأفراد في المحور الثانى				**1,00**

9. تمثيل العناصر الإضافية

يعتبر تمثيل العناصر الإضافية مهم جيدا عند القيام بتحليل المعطيات سواء كانت أفراد أو متغيرات ؛ حيث تسمح هذه العملية من إدخال في التحليل والتفسير المتغيرات أو الأفراد التي لم يكون لها دور في عملية الحساب، على سبيل المثال في الاستبيانات فإن المتغيرات مثل (الجنس، السن، الطبقة الاجتماعية ... الخ) يتم وضعها كعناصر إضافية من أجل إثراء التفسير و التحليل ؛ أما الأفراد يمكن وضعها كعناصر إضافية إذا لم يتم إدراجها لكون خصائصها تختلف بكثير عن الأفراد الأخرى التي تم إدخالها عند قيمنا بالتحليل، أو من أجل معرفة مكانة نقطة معينة بالنسبة للأفراد الأخرى أو مجموعة معينة.

1.9. مركبات المتغيرات الإضافية

يجب حساب موقع هذه المتغيرات بالنسبة لمركز ثقل O للسحابة المدروسة، وذلك من أجل إسقاط المتغيرات الجديدة بالنسبة لمتغيرات الأخرى، وبالتالي يتم حساب مركبات المتغيرة الإضافية s عن طريق حساب المعاملات المركزة لهذه الأخيرة، و يتم الحصول على إحداثيات النقاط s بضرب الشعاع $\vec{OS}$ بالشعاع الوحدوي V_α الحامل للمحور ذات الرتبة α، إحداثيات المتغيرات الإضافية هي مركبات الشعاع $X_{(p,n)}V_{\alpha(n,1)}$، حيث X يتم حسابها باستعمال المركبات المتغيرات s كالأتي : $X_{is}=\frac{X_{is}-\bar{X}_s}{S_s\sqrt{n}}$

أما مركبات شعاع V_α أي الأشعة الذاتية في $\Re^n$ للمتغيرات يمكن الحصول عليها باستعمال علاقة العبور التالية:

$$V_\alpha(i) = \lambda_\alpha^{-1/2} \sum_{j=1}^{p} X_{ij} A_\alpha(j)$$

مثال 12 : بأخذ المثال 11، لنفترض أننا لدينا المتغيرة الإضافية التالية x_4 :

	X4
A	150
B	250
C	240
D	7

يتم إسقاط المتغيرة X_4 عن طريق حساب الأشعة الذاتية V_α كما يلي :

أ- المحور الأول : لدينا الشعاع الذاتي a_1 :

a1	-0.54	0.56	0.63

والقيمة الذاتية في المحور الأول λ_1=2 ,12

إذا الأشعة الذاتية V_1 يتم الحصول عليها باستعمال الجدول المركز والمختصر التالي :

تحليل المعطيات

	x1	x2	x3
X	0,3237	-1,3416	-1,5076
	-0,2806	-0,4472	0,9045
	1,3596	0,4472	-0,3015
	-1,4028	1,3416	0,9045

وبقسمة هذا الأخير على $\sqrt{n}$ نحصل على :

x1	X2	x3
0,1618	-0,6708	-0,7537
-0,1402	-0,2236	0,4522
0,6798	0,2236	-0,1507
-0,7014	0,6708	0,4522

باستعمال العلاقة : $V_\alpha(i)=\lambda_\alpha^{-1/2}\sum_{j=1}^{p}X_{ij}A_\alpha(j)$

نحصل على الشعاع الذاتي V_1 :

V_1	-0,64
	0,16
	-0,23
	0,71

وبأخـذ الـشعاع الإضـافي x4 وبتحويلـه إلى قـيم ممركـزة ومختصرة، وبقسمته على $\sqrt{n}$ ، فان منقول هذا الأخير هـو :

x4	-0.0602	0.4527	0,4014	-0,7938

وبضرب الشعاع X4 بالقيم الذاتية V1 نحصل على إحداثيات المتغيرة X4 في المحور الأول :

المتغيرة	المحور الأول
X4	-0,54

بنفس الطريقة نستخرج مركبة المتغيرة X4 في المحور الثاني.

ب- المحور الثاني : لدينا الشعاع الذاتي a2 :

A2	0.76	0.64	0.07

والقيمة الذاتية المقابلة له في المحور الثاني $\lambda_2 = 0,61$

إذا الشعاع الذاتي V_2 يتم الحصول عليها باستعمال الجدول المركز والمختصر مع استعمال العلاقة التالية :

$$V_\alpha(i) = \lambda_\alpha^{-1/2} \sum_{j=1}^{p} X_{ij} A_\alpha(j)$$

نحصل على الشعاع الذاتي V_2 في R^4 :

V_2	-0,45
	-0,27
	0,83
	-0,09

وبأخذ الشعاع الإضافي x4 وبتحويله إلى قيم ممركزة ومختصرة، وبقسمته على $\sqrt{n}$ ، فان منقول هذا الأخير هو :

X4	-0.0602	0.4527	0,4014	-0,7938

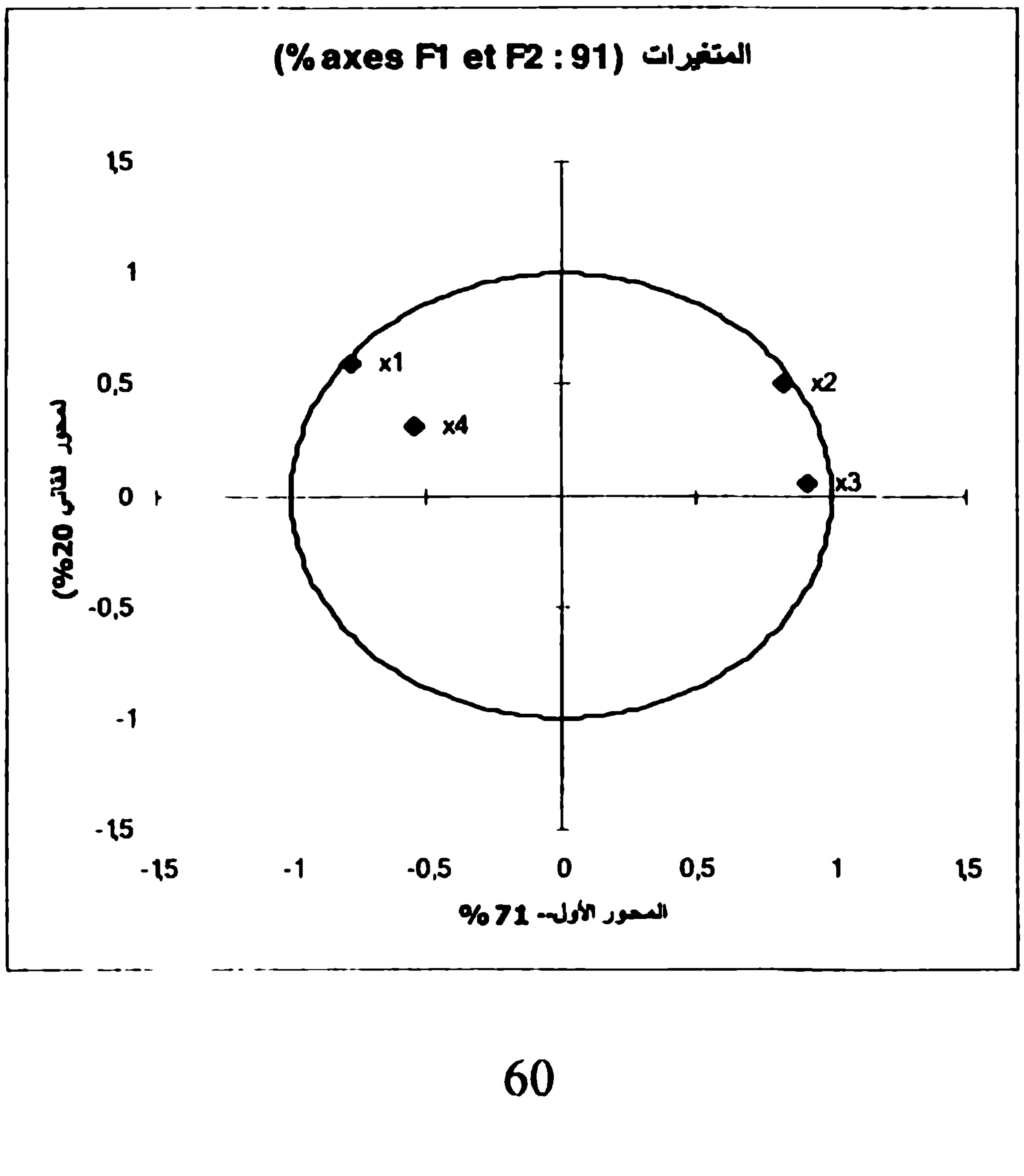
نحصل بضرب الشعاع X4 بالقيم الذاتية V2 على إحداثيات المتغيرة X4 في المحور الثاني :

المحور الثاني	المتغيرة
0,30	X4

وبتالي مركبات المتغيرة x4 في المحور الأول و الثاني هي :

المحور الثاني	المحور الأول	المتغيرة
0,30	-0,54	X4

نقوم بإسقاط هذه المتغيرة في المستوي السابق مع المتغيرات الأخرى نحصل على :



ملاحظة : لا يمكن حساب للمتغيرة الإضافية نسبة التمثيل النسبية ، ولكن يمكن أن نحسب لها نسبة التمثيل المطلقة والتي قيمها :

المتغيرة	Cos2(α1)	cos2(α2)	المستوي
X4	0,30	0,10	0,40

وعليه فان المتغيرة X4 غير ممثلة جيدا في هذا المستوي، وهذا ما نلاحظه في الشكل البياني حيث المتغيرة X4 تبتعد عن محيط الدائرة، عكس المتغيرات الأخرى، التي نلاحظ أنها تقترب من المحيط، وبالتالي فهي ممثلة جيدا في هذا المستوي ويمكن أن ندعمه حسابيا :

المتغيرة	Cos2 (α1)	Cos2 (α2)	المستوي
X1	0,61	0,35	0,96
X2	0,68	0,25	0,93
X3	0,84	0	0,84

2.9. مركبات الأفراد الإضافية

يتم إسقاط الفرد الجديد بالنسبة للأفراد الأخرى، عن طريق حساب موقع الفرد بالنسبة لمركز ثقل O للسحابة المدروسة ، وبذلك حساب مركبات الفرد الإضافي s عن طريق حساب المعاملات المركزة لهذه الأخيرة كما يلي :

$$X_{sj}=\frac{X_{sj}-\bar{X}_j}{S_j\sqrt{n}}$$

ويتم الحصول على إحداثيات النقاط s عن طريق ضرب الشعاع $\vec{OS}$ بالشعاع الوحدوي a_α الحامل للمحور ذات الرتبة α ؛ أي أن إحداثيات النقاط الأفراد الإضافية هي مركبات الشعاع $(x_{sj})a'_\alpha$.

مثال 13 : بأخذ المثال 11، لتكن النقطة الإضافية التالية E :

المتغيرات / الأفراد	X1	X2	X3
E	10	20	30

يتم تمثيل مكانة هذه النقطة في السحابة، باستعمال المعطيات المتعلقة بالمتوسط والانحراف المعياري للمتغيرات :

I	X1	X2	X3
المتوسط	46,25	6,5	2,25
الانحراف المعياري	11,5839	1,1180	0,8291

يتم تحويل الفرد E باستعمال العلاقة $X_{sj}=\frac{X_{sj}-\bar{X}_j}{S_j\sqrt{n}}$ أي :

I	X1	X2	X3
E	-1,56	6,03	16,73

وعليه إحداثيات الفرد E في المحور الأول و الثاني هي :

$$Y_t = X_{sj}'A = (-1,56 \quad 6,03 \quad 16,73)\begin{pmatrix} -0,54 & 0,76 \\ 0,76 & 0,56 \\ 0,36 & 0,63 \end{pmatrix} = (11,44 \quad 12,73)$$

المحور الثاني	المحور الأول	الفرد
12,73	11,44	E

تمارين :

1.1 : لدينا المعطيات التالية المتعلقة بمنتوج معين كمايلي :

الأفراد	المبيعات X3	نفقات الاشهار X2	الأسعار X1
الجزائر العاصمة	70	15	30
وهران	55	20	14
عنابة	40	17	25
ورقلة	30	14	31

1. احسب متوسط كل متغيرة، ثم حول هذا الجدول إلى جدول ممركز.

2. اذا رمزنا للجدول الممركز بـ X، احسب $\frac{1}{4}X'X$.

3. القيم الذاتية لمصفوفة التباين والتباين المشترك أي $\frac{1}{4}X'X$ هي:
$\lambda_1 = 213.99$ ، $\lambda_3 = 0.11$ استنتج القيم الذاتية الباقية : $\lambda_2 = ?$.

4. احسب نسبة تمثيل للمستوي.

5. مع علم الأشعة الذاتية المقابلة للقيم الذاتية العظمى هي :
$a_1 = \begin{bmatrix} -0.10 \\ 0.04 \\ 0.99 \end{bmatrix}$ ، $a_2 = \begin{bmatrix} 0.94 \\ -0.31 \\ 0.11 \end{bmatrix}$ ، مثل بيانيا الأفراد.

2.1 : بأخذ معطيات التمرين 1

1. أوجد كل من نسبة تمثيل الأفراد في المستوي $\cos^2(\alpha)$ ، ثم استنتج نسبة تمثيل في المحور الثالث.

2. أوجد مساهمة الأفراد في كل محور ؛ مع العلم أن الشعاع الذاتي المقابل لقيمة الذاتية الثالثة هو $a_3 = \begin{bmatrix} 0.32 \\ 0.95 \\ -0.01 \end{bmatrix}$

3. مثل بيانيا المتغيرات في المستوي.

4. اشرح النتائج المحصل عليها في كل من التمرين 1 و 2.

3.1 : لدينا الجدول التالي المتعلق بالمتغيرات الاقتصادية للدول العربية :

الأفراد (الدول)	X3	X2	X1
الجزائر	70	15	1000
تونس	55	20	700
مصر	40	17	2200
ليبيا	30	14	1500

ملاحظة : المقياس الذي يجب أخذه هو $D_{\frac{1}{\sigma}}$.

1. احسب متوسط كل متغيرة وانحرافها المعياري، ثم حول هذا الجدول إلى جدول ممركز.

2. أحسب مصفوفة الارتباطات.

3. القيم الذاتية لمصفوفة الارتباطات هي : $\lambda_1 = 1.82$ ، $\lambda_2 = 0.82$ ، استنتج القيم الذاتية الباقية : ? = λ_3

4. احسب نسبة التمثيل للمستوي.

5. مع علم الأشعة الذاتية المقابلة للقيم الذاتية العظمى هي :

$a_1 = \begin{bmatrix} 0.65 \\ -0.43 \\ -0.62 \end{bmatrix}$ ، $a_2 = \begin{bmatrix} 0.21 \\ 0.89 \\ -0.40 \end{bmatrix}$ ، مثل بيانيا الأفراد.

4.1 : بأخذ معطيات التمرين 1.

أوجد كل من نسبة تمثيل الأفراد في المستوي $\cos^2(\alpha)$

أوجد مساهمة الأفراد في كل محور.

مثل بيانيا المتغيرات.

اشرح النتائج المحصل عليها في كل من التمرين 3 و 4.

5.1 : طبق على جدول التمرين 2 نفس المراحل المذكورة ؛ باستعمال مقياس $D_{\frac{1}{\sigma}}$.

الفصل الثاني

تحليل المعطيات بالمعاملات للتوفيقات

Analyse Factorielle
des Correspondances AFC

تمهيد

يعود استعمال طريقة AFC إلى الفرنسي Jean-Paul Benzcri في 1961، حيث قام بدراسة العلاقة مابين متغيرتين كيفيتين باستعمال المركبات؛ تتمثل الطريقة رياضيا في تطبيق طريقة ACP مع استعمال وحدة قياس كاي تربيع ($\chi 2$)، وبصفة عامة تستعمل طريقة أو تقنية AFC في جدول مزدوج ذات متغيرات كيفية، حيث أنها تسمح من تمثيل نقاط الأسطر ونقاط الأعمدة لجدول المزدوج ذات متغيرات كيفية.

2. عرض جدول المعطيات

بمأن الجدول الذي نطبق عليه AFC هو جدول ذات متغيرات كيفية، فان تقاطع كل عمود وسطر؛ يعطي تكرار أي عدد الأفراد ذات خاصية تقاطع السطر و العمود، وبالتالي فان محتوى الجدول عبارة عن تكرارات، أي أعداد طبيعية.

$Z_{n\times p} =$

X/ Y	1 2 ... j ... p
1	n_{11} n_{12} . .. n_{1j} ... n_{1p}
2	.
.	n_{i1}. . . n_{ij} ..... n_{ip}
i	.
.	
n	n_{n1}. . . n_{nj} ..... n_{np}

تتمثل طريقة AFC في تحليل سحابة الأسطر N(I) وسحابة الأعمدة N(J).

مثال 14 : قامت مؤسسة بدراسة انطباع المستهلكين فيما يخص منتوجها في ست ولايات من الوطن ؛ فتحصلنا على ما يلي :

الولايات / انطباع المستهلك فيما يخص المنتوج	الجزائر	وهران	عنابة	ورقلة	المجموع
ممتاز	9	35	24	8	76
جيد	66	72	112	48	298
متوسط	52	86	43	26	207
رديء	55	103	40	25	223
المجموع	182	296	219	107	804

3. الجدول الاستكشافي للجدول المزدوج

من أجل تطبيق تحليل المعطيات بالمركبات الرئيسية المتمثلة في توضيح الارتباطات للجداول المزدوجة (جداول الاقتران) ؛ يجب إيجاد مجموعة من الجداول التي تمكننا من توضيح ذلك.

1.3. جدول التكرارات النسبية

لنعتبر الجدول السابق $Z_{n\times p}$ ذات النمط $n\times p$ حيث $(n\geq p)$، وبوضع :

$$n_{..}=\sum_{i=1}^{n}\sum_{j=1}^{p}n_{ij} \quad ، \quad n_{.j}=\sum_{i=1}^{n}n_{ij} \quad و \quad n_{i.}=\sum_{j=1}^{p}n_{ij}$$

يــتم المــرور مــن جــدول التكــرارات إلى جــدول التكــرارات النــسبية؛ باســتعمال العلاقــة التاليــة : $f_{ij} = {n_{ij}}/{n_{..}}$ ، حيــث تعطــى التكرارات النسبية للهوامش كما يلي : $f_{i.} = {n_{i.}}/{n_{..}}$ و $f_{.j} = {n_{.j}}/{n_{..}}$.

يعطى جدول التكرارات النسبية كما يلي :

		J			
	1 . . .	j . . .	p		
I	1	f_{11} . . .	f_{1j} . . .	f_{1p}	$f_{1.}$
	.	.	.	.	.
	i	f_{11} .	$f_{ij} = n_{ij}/n_{..}$.	f_{ip}	$f_{i.}$
	.	.	.	.	.
	n	f_{n1} . . .	f_{nj} . . .	f_{np}	$f_{n.}$
		$f_{.1}$.	$f_{.j}$.	$f_{.p}$	1

احتمالات لكل من $\Pr ob(X=i)$ و $\Pr ob(Y=j)$ مقدر بـ :

$$\hat{\Pr} ob(X=i) = f_{i.} \quad \text{و} \quad \hat{\Pr} ob(Y=j) = f_{.j}$$

1.1.3. جدول التكرارات النسبية للأسطر : من أجل مقارنة الأسـطر فيمـا بينهـا يجـب بنـاء جـدول النـسبي للأسـطر، والـذي يـتم استخراجه باستعمال العلاقة التاليـة :

$$f_j^i = \frac{n_{ij}}{n_{i.}} = \frac{f_{ij}}{f_{i.}}$$

تمثل الكمية f_j^i التقدير التالي $\widehat{\Pr ob}(Y = j \setminus X = i)$ ؛ ومنه يعطى جدول التكرارات النسبية للأسطر كما يلي :

$$
XI_{n\times p} =
\begin{array}{c}
\\ \\ 1 \\ . \\ . \\ . \\ i \\ . \\ . \\ . \\ n \\ \\
\end{array}
\begin{array}{c}
J \\
1 \quad . \quad . \quad . \quad j \quad . \quad . \quad . \quad p \\
\left[\begin{array}{ccccccc}
n_{11}/n_{1.} & . \; . \; . & n_{1j}/n_{1.} & . \; . \; . & n_{1p}/n_{1.} \\
. & & . & & \\
. & & & & \\
. & & . & & . \\
n_{i1}/n_{i.} & . & n_{ij}/n_{i.} & . \; . \; . & n_{ip}/n_{i.} \\
. & & . & & . \\
. & & . & & \\
. & & . & & . \\
n_{n1}/n_{n.} & . \; . \; . & n_{nj}/n_{n.} & . \; . \; . & n_{np}/n_{n.}
\end{array}\right] \\
f_{.1} \quad . \qquad f_{.j} \qquad . \quad f_{.p}
\end{array}
\begin{array}{c}
\\ \\ 1 \\ . \\ . \\ . \\ 1 \\ . \\ . \\ . \\ 1 \\ p \\
\end{array}
$$

مثال 15 : بأخذ المثال 14 فإن التكرارات النسبية للأسطر (توزيع انطباع المستهلك حسب الولايات التي ينتمي إليها) :

الولايات / انطباع المستهلك فيما يخص المنتوج	الجزائر	وهران	عنابة	ورقلة
ممتاز	9/76	35/76	24/76	8/76
جيد	66/298	72/298	112/298	48/298
متوسط	52/207	86/207	43/207	26/207
رديء	55/223	103/223	40/223	25/223
الهامش	182/804	296/804	219/804	107/804

2.1.3. جدول التكرارات النسبية للأعمدة : يجب بناء جدول النسبي للأعمدة من أجل مقارنة الأعمدة فيما بينها ، ويتم استنتاجه باستعمال العلاقة التالية :

$$f_i^j = \frac{n_{ij}}{n_{.j}} = \frac{f_{ij}}{f_{.j}}$$

تمثل الكمية f_i^j تقدير الاحتمال الشرطي $\hat{\Pr}ob(X = i \setminus Y = j)$ ؛ يعطى جدول التكرارات النسبية للأعمدة كما يلي :

$$
{}^{t}XJ_{pxn}=
\begin{array}{c}
 \\
 \\
1\\ \vdots \\ j \\ \vdots \\ p \\
 \\
\end{array}
\begin{array}{c}
I \\
1 \quad . \quad . \quad . \quad i \quad . \quad . \quad . \quad n \\
\left[\begin{array}{ccccc}
n_{11}/n_{.1} & \cdots & n_{i1}/n_{.1} & \cdots & n_{n1}/n_{.1} \\
\vdots & & \vdots & & \vdots \\
n_{1j}/n_{.j} & \cdots & n_{ij}/n_{.j} & \cdots & n_{nj}/n_{.j} \\
\vdots & & \vdots & & \vdots \\
n_{1p}/n_{.p} & \cdots & n_{ip}/n_{.p} & \cdots & n_{np}/n_{.p}
\end{array}\right] \\
f_{1.} \quad . \quad . \quad . \quad f_{i.} \quad . \quad . \quad . \quad f_{n.} \quad n
\end{array}
\begin{array}{c}
 \\
 \\
1\\ \vdots \\ 1 \\ \vdots \\ 1 \\
 \\
\end{array}
$$

مثال 16 : نأخذ المثال 14 فإن التكرارات النسبية للأعمدة (توزيع الولايات حسب انطباع المستهلك في المنتوج).

انطباع المستهلك فيما يخص المنتوج/ الولايات	ممتاز	جيد	متوسط	رديء
الجزائر	9/182	66/182	52/182	55/182
وهران	35/296	72/296	86/196	103/296
عنابة	24/219	112/219	43/219	40/219
ورقلة	8/107	48/107	26/107	25/107
الهامش	76/804	289/804	207/804	223/804

أ- مؤشر التجاذب/التنافر : يعرف مؤشر التجاذب / التنافر dij بـ :

$$d_{ij} = \frac{f_{ij}}{f_{i.}f_{.j}}$$

واستعملت هذه العلاقة من طرف كل من Escofier & Pagès (1988) وتعلب هذه الأخيرة دور أساسي في AFC.

- إذا كان : dij > 1 نقول عن الخاصية i و j تتجاذب في الحالة المعاكس نقول عنها تتنافر.

ويمكن كتابتها كمايلي :

$$d_{ij} = \frac{f_{ij}}{f_{i.}f_{.j}} = \frac{\frac{f_{ij}}{f_{i.}}}{f_{.j}} = \frac{\hat{\Pr}ob(Y = j \backslash X = i)}{\hat{\Pr}ob(Y = j)}$$

أو :

$$d_{ij} = \frac{f_{ij}}{f_{i.}f_{.j}} = \frac{\frac{f_{ij}}{f_{.j}}}{f_{i.}} = \frac{\hat{\Pr}ob(X = i \backslash Y = j)}{\hat{\Pr}ob(X = i)}$$

ملاحظة :

1. يمكن القول عن الخاصية [Y=j] أنها ممثلة بصفة معتبرة (أكثر تمثيلا) في المجتمع [X=i] مقارنة بالمجتمع الكلي، نفس الشيء بالنسبة للخاصية [X=i] أنها ممثلة بصفة معتبرة في المجتمع [Y=j] مقارنة بالمجتمع الكلي في حالة ما إذا كانت dij تفوق الواحدة، وفي الحالة المعاكسة نقول أنها أقل تمثيلا.

2. إذا كان dij=1 نقول أن لا يوجد ارتباط بين السطر i والعمود j.

مثال 17 : بأخذ المثال 14 مؤشر التقارب/التنافر معطى في الجدول التالي :

الولايات / انطباع المستهلك فيما يخص المنتوج	الجزائر	وهران	عنابة	ورقلة
ممتاز	0,52	1,25	1,15	0,79
جيد	0,97	0,65	1,37	1,21
متوسط	1,10	1,12	0,76	0,94
ردىء	1,08	1,25	0,65	0,84

$$d_{ij} = \frac{f_{ij}}{f_{i.}f_{.j}} = \frac{n_{ij}/n}{(n_{i.}/n_{..})(n_{.j}/n_{..})} = \frac{n_{ij}/n_{.j}}{n_{i.}/n_{..}}$$

$$d_{23} = \frac{f_{23}}{f_{2.}f_{.3}} = \frac{n_{23}/n_{..}}{(n_{2.}/n_{..})(n_{.3}/n_{..})} = \frac{n_{23}/n_{.3}}{n_{2.}/n_{..}} = \frac{112/219}{298/804} = 1.37$$

ب- اختبار كاي تربيع (كاي $\chi 2$)

عند دراسة ظاهرة كيفية متعلقة بمتغيرتين كيفيتين، يمكن البحث عن وجود أو عدم وجود علاقة بين هاتين المتغيرتين ؛ وذلك باستعمال اختبار كاي تربيع المعطى بالعلاقة التالية :

$$X^2 = \sum_{i=1}^{n} \sum_{j=1}^{p} \frac{(n_{ij} - (n_{i.} n_{.j} / n_{..}))^2}{n_{i.} n_{.j} / n_{..}}$$

بالنسبة للمثال 14 قيمة هذه الإحصائية هي : 51,73.

والتي تم حسابها من الجدول التالـي :

جدول مساهمة كل خانة في كاي تربيع

الولايات / انطباع المستهلك فيما يخص المنتوج .	الجزائر	وهران	عنابة	ورقلة	المجموع
ممتاز	3,91	1,76	0,53	0,44	6,64
جيد	0,03	12,96	11,71	1,75	26,46
متوسط	0,56	1,26	3,18	0,09	5,09
رديء	0,40	5,32	7,08	0,74	13,55
المجموع	4,91	21,30	22,49	3,02	51,73

بمقارنتها مع القيمة الحرجة (المجدولة) $\chi^2_{0.95}$ بدرجة حرية (p-1) (n-1). في هذا المثال درجة الحرية هي : 9 وبالتالي قيمة كاي تربيع المجدولة هي : 16,9 أنظر الملحق 4 ؛ القيمة المحسوبة أكبر من القيمة المجدولة، منه فإنه يوجد ارتباط بين الأسطر والأعمدة، أي بين الولايات وانطباع المستهلك.

ملاحظة : ان تحليل المعطيات بالمعاملات AFC ؛ يمكن أن يعطي معلومات مهمة حتى و إن كان $\chi 2$ غير معنوي.

4. تحليل جدول محول (الأسطر أو الأعمدة)

1.4. تحليل جدول التكراري النسبي

1.1.4. تحليل جدول الأسطر : تحليل جدول الأسطر XI المرتبط بالسحابة N(I) كما يلي :

$$N(I)=\left\{\left(XI_i,\frac{n_{i\bullet}}{n_{\bullet\bullet}}\right),i=1,\ldots,n\right\}\subset R^p$$

حيث : $\frac{n_{i\bullet}}{n_{\bullet\bullet}}$ تمثل ثقل كل سطر في السحابة الأسطر.

2.1.4. تحليل جدول الأعمدة : تحليل جدول الأعمدة XJ مرتبط بالسحابة N(J) كما يلي :

$$N(J)=\left\{\left(XJ_j,\frac{n_{\bullet j}}{n_{\bullet\bullet}}\right),j=1,\ldots,p\right\}\subset R^n$$

حيث : $\frac{n_{\bullet j}}{n_{\bullet\bullet}}$ تمثل ثقل كل عمود في سحابة الأعمدة.

5. اختيار المقياس (Métrique)

1.5. مقياس نقاط الأسطر في فضاء Rp : إن مقياس المختار في فضاء الأسطر هو :

$$MI_{P\times p}=\begin{pmatrix} \frac{n_{\bullet\bullet}}{n_{\bullet 1}} & 0 & & & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \frac{n_{\bullet\bullet}}{n_{\bullet j}} & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & & & 0 & \frac{n_{\bullet\bullet}}{n_{\bullet p}} \end{pmatrix}$$

يمكن حساب البعد أو المسافة بين نقطتين XI_i و $XI_{i'}$ في جدول الأسطر كمايلي :

$$d^2(XI_i,XI_{i'})=\|XI_i-XI_{i'}\|^2_{MI}={}^t(XI_i-XI_{i'})MI(XI_i-XI_{i'})$$

$$d^2(XI_i,XI_{i'})=\sum_{j=1}^{q}\frac{n_{\bullet\bullet}}{n_{\bullet j}}\left(\frac{n_{ij}}{n_{i\bullet}}-\frac{n_{i'j}}{n_{i'\bullet}}\right)^2$$

2.5. **مقياس نقاط الأعمدة في فضاء R^n** : إن مقياس المختار في فضاء الأعمدة هو :

$$MJ_{n\times n}=\begin{pmatrix} \frac{n_{\bullet\bullet}}{n_{1\bullet}} & 0 & & & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \frac{n_{\bullet\bullet}}{n_{i\bullet}} & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & & & 0 & \frac{n_{\bullet\bullet}}{n_{n\bullet}} \end{pmatrix}$$

يمكن حساب البعد أو المسافة بين نقطتين $XJ_{j'}$ و XJ_j في جدول الأعمدة كما يلي :

$$d^2\left(XJ_j, XJ_{j'}\right) = \sum_{i=1}^{n} \frac{n_{\bullet\bullet}}{n_{i\bullet}} \left(\frac{n_{ij}}{n_{\bullet j}} - \frac{n_{ij'}}{n_{\bullet j'}} \right)^2$$

6. مركز الثقل

1.6. مركز الثقل لسحابة الأسطر

$$\underline{g}_I = \sum_{i=1}^{n} \frac{n_{i\bullet}}{n_{\bullet\bullet}} \underline{XI}_i$$

تكتب المركبة j من هذا الشعاع كمايلي :

$$g_I^j = \sum_{i=1}^{n} \frac{n_{i\bullet}}{n_{\bullet\bullet}} \frac{n_{ij}}{n_{i\bullet}} = \frac{n_{\bullet j}}{n_{\bullet\bullet}}$$

إذا شعاع مركز الثقل لسحابة الأسطر هو :

$$g_I = \begin{pmatrix} \frac{n_{\bullet 1}}{n_{\bullet\bullet}} \\ \vdots \\ \frac{n_{\bullet p}}{n_{\bullet\bullet}} \end{pmatrix}$$

2.6. مركز الثقل لسحابة الأعمدة

$$\underline{g}_J = \sum_{j=1}^{p} \frac{n_{\bullet j}}{n_{\bullet\bullet}} \underline{XJ}_j$$

تكتب المركبة i من هذا الشعاع :

$$g_J^i = \sum_{j=1}^{p} \frac{n_{\bullet j}}{n_{\bullet\bullet}} \frac{n_{ij}}{n_{\bullet j}} = \frac{n_{i\bullet}}{n_{\bullet\bullet}}$$

إذا شعاع مركز الثقل لسحابة الأعمدة هو :

$$g_J = \begin{pmatrix} \frac{n_{1\bullet}}{n_{\bullet\bullet}} \\ \vdots \\ \frac{n_{n\bullet}}{n_{\bullet\bullet}} \end{pmatrix}$$

7. مصفوفة التباينات

1.7. مصفوفة التباينات (Matrice d'inertie) لسحابة الأسطر :N(I)

نضع DI مصفوفة الثقل المرتبطة بنقاط الأسطر للسحابة N(I)

$$DI_{n\times n} = \begin{pmatrix} \frac{n_{1\bullet}}{n_{\bullet\bullet}} & 0 & & & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \frac{n_{i\bullet}}{n_{\bullet\bullet}} & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & & & 0 & \frac{n_{n\bullet}}{n_{\bullet\bullet}} \end{pmatrix}$$

حيث مصفوفة التباين للسحابة N (I) هي مصفوفة VI :

$$VI_{p\times p} = {}^t\widetilde{X}I_{p\times n} DI_{n\times n} \widetilde{X}I_{n\times p}$$

مع : ${}^t\widetilde{X}I = {}^t(XI - g_I)$

2.7. مصفوفة التباينات (Matrice d'inertie) لسحابة الأعمدة :N(J)

نضع DJ مصفوفة الثقل المرتبطة بنقاط الأعمدة للسحابة N(J).

$$DJ_{p\times p}=\begin{pmatrix}\frac{n_{\bullet 1}}{n_{\bullet\bullet}} & 0 & & & 0\\ 0 & \ddots & \ddots & & \vdots\\ \vdots & \ddots & \frac{n_{\bullet j}}{n_{\bullet\bullet}} & \ddots & \vdots\\ \vdots & & \ddots & \ddots & 0\\ 0 & \cdots & \cdots & 0 & \frac{n_{\bullet p}}{n_{\bullet\bullet}}\end{pmatrix}$$

حيث مصفوفة التباين للسحابة N(J) هي مصفوفة VJ :

$$VJ_{n\times n}={}^{t}\widetilde{X}J_{n\times p}\,DJ_{p\times p}\,\widetilde{X}I_{p\times n}$$

مع :

$${}^{t}\widetilde{X}J={}^{t}(XJ-g_{J})$$

8. المحاور الأساسية

1.8. المحاور الأساسية لساحبة الأسطر N(I)

نقوم بتقطير المصفوفة C_I من أجل إيجاد المحاور، والتي تكتب كما يلي:

$$C_I = MI^{1/2}\,VI\;\;MI^{1/2}$$

حيث يتمثل التقطير في تطبيق ACP على الجدول XI ذات المقياس MI المذكور سابقا المتعلق بالسطر، وليس المصفوفة الأحادية.

هذه المصفوفة حدها العام هو :

$$C_{jj'}=\frac{n_{\bullet\bullet}}{\sqrt{n_{\bullet j}n_{\bullet j'}}}V_{jj'}=\frac{n_{\bullet\bullet}}{\sqrt{n_{\bullet j}n_{\bullet j'}}}\left(\left(\sum_{i}\frac{n_{ij}n_{ij'}}{n_{\bullet\bullet}n_{i\bullet}}\right)-\frac{n_{\bullet j'}n_{\bullet j}}{n_{\bullet\bullet}^{2}}\right)$$

$$C_{jj'} = \sum_i \frac{n_{ij} n_{ij'}}{n_{i\bullet} \sqrt{n_{\bullet j} n_{\bullet j'}}} - \frac{\sqrt{n_{\bullet j} n_{\bullet j'}}}{n_{\bullet\bullet}}$$

هذه المصفوفة CI يمكن كتابتها بدلالة المصفوفة H ، والتي حدها يمكن كتبته كما يلي :

$$h_{ij} = \frac{n_{ij}}{\sqrt{n_{i\bullet} n_{\bullet j}}} - \frac{\sqrt{n_{i\bullet} n_{\bullet j}}}{n_{\bullet\bullet}}$$

وعليه فإن CI تكتب كما يلي :

$$C_I = {}^t HH$$

2.8. المحاور الأساسية لساحبة الأعمدة N(J)

من أجل إيجاد المحاور نقوم بتقطير المصفوفة التالية :

$$C_J = MJ^{1/2}\ VJ\ \ MJ^{1/2}$$

وهي تتمثل في تطبيق ACP على الجدول XJ ذات المقياس MJ المذكور سابقا المتعلق بالأعمدة، و ليس المصفوفة الأحادية.

هذه المصفوفة حدها العام يعطى كما يلي :

$$b_{jj'} = \sum_i \frac{n_{ji} n_{j'i}}{n_{\bullet i} \sqrt{n_{j\bullet} n_{j'\bullet}}} - \frac{\sqrt{n_{j\bullet} n_{j'\bullet}}}{n_{\bullet\bullet}}$$

يمكن حسابها عن طريق المصفوفة H المذكور سابقا كما يلي :

$$C_J = H{}^t H$$

3.8. مصفوفة المحاور الأساسية للأسطر والأعمدة

كل من المصفوفة CI و المصفوفة CJ لهما نفس القيم الذاتية ، إذا كانت λ قيمة ذاتية موجبة ؛ فإن استخراج هذه الأخيرة والأشعة المقابلة لها يتم عن طريق ما يلي :

$$C_I a_I = \lambda a_I \quad \wedge \quad {}^t a_I a_I = 1$$

$$C_J a_J = \lambda a_J \quad \wedge \quad {}^t a_J a_J = 1$$

1.3.8. الأشعة الذاتية: يكتب الحد العام للأشعة الذاتية للأسطر والأعمدة على الترتيب كمايلي :

$$a_J = \pm\frac{1}{\sqrt{\lambda}} H a_I \quad , \quad a_I = \pm\frac{1}{\sqrt{\lambda}} {}^t H a_J$$

كل عنصر من الشعاع الأسطر أو الأعمدة يكتب على الشكل التالي :

$$\forall j : a_I^j = \pm\frac{1}{\sqrt{\lambda n_{\bullet j}}} \sum_i \frac{n_{ij}}{n_{i\bullet}} a_J^i \ , \ \forall i : a_J^i = \pm\frac{1}{\sqrt{\lambda n_{i\bullet}}} \sum_j \frac{n_{ij}}{n_{\bullet j}} a_I^j$$

يمكن البرهان على صحة هذه العلاقات.

ملاحظة :

1. الشعاع $\begin{pmatrix} \sqrt{n_{\bullet 1}} \\ \vdots \\ \sqrt{n_{\bullet p}} \end{pmatrix}$ هو الشعاع الذاتي للمصفوفة C_I المرافقة للقيمة الذاتية 0.

2. الشعاع $\begin{pmatrix}\sqrt{n_{1\bullet}}\\ \vdots \\ \sqrt{n_{n\bullet}}\end{pmatrix}$ هو الشعاع الذاتي للمصفوفة C_I المرافقة للقيمة الذاتية 0.

3. المصفوفة C_I يمكن تفكيكها إلى مصفوفتين ؛ المصفوفة C حدها العام $a_{jj'}$ ، والمصفوفة T حدها العام $d_{jj'}$

أي : $C_I = C - T$ ، والتي تكتب عن طريق حدها العام كما يلي :

$$C_{jj'} = \underbrace{\sum_i \frac{n_{ij} n_{ij'}}{n_{i\bullet}\sqrt{n_{\bullet j} n_{\bullet j'}}}}_{a_{jj'}} - \underbrace{\frac{\sqrt{n_{\bullet j} n_{\bullet j'}}}{n_{\bullet\bullet}}}_{d_{jj'}}$$

أ- كل شعاع ذاتي لـ C_I له قيمة ذاتية $\lambda>0$ هو قيمة ذاتية لـ C المرافقة بالقيمة الذاتية λ.

ب- الشعاع $\begin{pmatrix}\sqrt{n_{\bullet 1}}\\ \vdots \\ \sqrt{n_{\bullet p}}\end{pmatrix}$ هو الشعاع الذاتي لـ C_I المرافق للقيمة الذاتية 0، وهو كذلك شعاع ذاتي لـ C المرافق للقيمة الذاتية 1.

مثال 18: نأخذ المثال 17

الولايات / انطباع المستهلك فيما يخص المنتوج	الجزائر	وهران	عنابة	ورقلة	المجموع
ممتاز	9	35	24	8	76
جيد	66	72	112	48	298
متوسط	52	86	43	26	207
رديء	55	103	40	25	223
المجموع	182	296	219	107	804

- أحسب القيم الذاتية و الأشعة الذاتية لجدول الأسطر.

الحل :

أولا : يجب حساب مصفوفة الأسطر CI، والتي يتم حسابها من خلال المصفوفة H، ومن أجل تبسيط الحساب نقوم بحساب المصفوفة C، وذلك من خلال المصفوفة :

$$h_{ij} = \frac{n_{ij}}{\sqrt{n_{i\bullet} n_{\bullet j}}},$$

بدلا من المصفوفة :

$$h_{ij} = \frac{n_{ij}}{\sqrt{n_{i\bullet} n_{\bullet j}}} - \frac{\sqrt{n_{i\bullet} n_{\bullet j}}}{n_{\bullet\bullet}}$$

إذا المصفوفة H للجدول السابق هي :

H

9/√(182*76)	35/√ (296*76)	24/√ (219*76)	8/√ (76*107)
66/√ (182*298)	72/√ (296*298)	112/√ (219*298)	48/√107*298)
52/√ (182*207)	86/√ (296*207)	43/√ (219*207)	26/√ (107*207)
55/√ (182*223)	103/√ (296*223)	40/√ (219*223)	25/√ (107*223)

المصفوفة C هي :

C=ᵗHH

C

0,23	0,29	0,24	0,17
0,29	0,39	0,29	0,21
0,24	0,29	0,30	0,20
0,17	0,21	0,20	0,14

يتم حساب القيمة الذاتية لهذه المصفوفة عن طريق المحدد $|C-\lambda I|$ ، فنحصل على :

CI	C	القيمة الذاتية
تعادل القيمة الذاتية 0	قيمة الذاتية 1	1
تعادل القيمة الذاتية الأولى λ1	القيمة الذاتية الأولى λ1	0,0552
تعادل القيمة الذاتية الثانية λ2	القيمة الذاتية الثانيةλ2	0,0090
تعادل القيمة الذاتية الثالثة λ3	القيمة الذاتية الثالثة λ3	0,000009

نسبة التمثيل للمستوي هي :

$$I = \frac{\lambda_1 + \lambda_2}{traceC - 1} \approx 100\%$$

الأشعة الذاتية للمصفوفة C هي :

a	a 1	A 2	a 3
13,49	0,09-	0,79	0,38-
17,20	0,67-	0,42-	0,04
14,79	0,69	0,38-	0,31-
10,34	0,24	0,23	0,87

2.3.8. المحاور الرئيسية : نستعمل لحساب المحاور الرئيسية نفس طريقة ACP حيث :

أ- بالنسبة لسحابة N(I) الأشعة $u_{I1}, u_{I2}, \ldots, u_{Iq}$ هي :

$$u_{Ik} = M_I^{\frac{1}{2}} a_{Ik} \Leftrightarrow u_{Ik}^{j} = \sqrt{\frac{n_{\bullet j}}{n_{\bullet\bullet}}} a_{Ik}^{j}, \forall j$$

ب- بالنسبة لسحابة N(J) الأشعة $u_{J1}, u_{J2}, \ldots, u_{Jq}$ هي :

$$u_{Jk} = M_J^{-\frac{1}{2}} a_{Ik} \Leftrightarrow u_{Jk}^{i} = \sqrt{\frac{n_{i\bullet}}{n_{\bullet\bullet}}} a_{Jk}^{i}, \forall i$$

9. المركبات الرئيسية

1.9. المركبات الرئيسية لسحابة الأسطر N(I): تتمثل المركبات الرئيسية لهذه السحابة في الشعاع y_I^k ، المحصل عليه عن طريق ACP على سحابة السطر وبالتالي علاقة المركبات تعطى كما يلي :

$$y_l^k = {}^tX_l M_l u_{lk} = {}^tX_l M_l^{\frac{1}{2}} a_{lk}$$

$$y_{li}^k = \sum_j X_{li}^j \sqrt{\frac{n_{\bullet\bullet}}{n_{\bullet j}}} a_{lk}^j$$

$$\boxed{y_{li}^k = \sum_j \frac{n_{ij}}{n_{i\bullet}} \sqrt{\frac{n_{\bullet\bullet}}{n_{\bullet j}}} a_{lk}^j}$$

مثال 19 : بأخذ المثال 18

الولايات / انطباع المستهلك فيما يخص المنتوج	الجزائر	وهران	عنابة	ورقلة	المجموع
ممتاز	9	35	24	8	76
جيد	66	72	112	48	298
متوسط	52	86	43	26	207
ردىء	55	103	40	25	223
المجموع	182	296	219	107	804

تحسب المركبات الرئيسية للأسطر في المحور الأول باستعمال العلاقة التالية :

$$y_{li}^1 = \sum_j \frac{n_{ij}}{n_{i\bullet}} \sqrt{\frac{n_{\bullet\bullet}}{n_{\bullet j}}} a_{lk}^j$$

a 1	- 0,09	- 0,67	0,69	0,24

ممتاز	y^1_{I1}	9/76 *RACINE((804/182))* (-0,09)+(35/76)*RACINE((804/296))* (-0,67)+(24/76)*RACINE((804/219))*0,69 +(8/76)*RACINE((804/107))*0,24	= -0,0419

الأسطر	y^1_{Ii}	y^1_{Ii}
ممتاز	y^1_{I1}	-0,0419
جيد	y^1_{I2}	0,2972
متوسط	y^1_{I3}	-0,1481
رديء	y^1_{I4}	-0,2453

نحـصل علـى المركبـات الرئيـسية للأسـطر في المحـور الثـاني ؛ باستعمال نفس الطريقة السابقة مع تغيير الشعاع الذاتي :

a 2
0,79
0,42-
0,38-
0,23

ممتاز	(9/76)*RACINE((804/182))*0,79+(35/76)*RACINE((804/296))*(-0,42)+(24/76)*RACINE((804/219))*(-0,38)+(8/76)*RACINE((804/107))*0,23	-0,2926

نحصل على :

الأسطر	y^2_{Ii}	y^2_{Ii}
ممتاز	y^2_{I1}	-0,2926
جيد	y^2_{I2}	0,0216
متوسط	y^2_{I3}	0,0510
رديء	y^2_{I4}	0,0235

غير أنه يمكن حساب المركبات باستعمال العلاقة التالية :

$$y^k_{Ii} = \sqrt{\lambda_k} \sqrt{\frac{n_{\bullet\bullet}}{n_{i\bullet}}} a^i_{jk}$$

ولكن تتطلب منا هذه العلاقة حساب الأشعة الذاتية للمصفوفة الأعمدة.

2.9. المركبات الرئيسية لسحابة الأعمدة N(J) : تتمثل المركبات الرئيسية لهذه السحابة في الشعاع y^k_J ، والذي تم الحصول عليه عن طريق تطبيق ACP على سحابة الأعمدة كما يلي :

$$y^k_J = {}^t X_J M_J u_{Jk} = {}^t X_J M_J^{\frac{1}{2}} a_{Jk}$$

$$y^k_{Jj} = \sum_i X'_{Jj} \sqrt{\frac{n_{\bullet\bullet}}{n_{i\bullet}}} a^i_{Jk}$$

أي أن :

$$y_{Jj}^{k}=\sum_{i}\frac{n_{ij}}{n_{\bullet j}}\sqrt{\frac{n_{\bullet\bullet}}{n_{i\bullet}}}a'_{jk}$$

أو بصياغة أخرى :

$$\boxed{y_{Jj}^{k}=\sqrt{\lambda_k}\sqrt{\frac{n_{\bullet\bullet}}{n_{\bullet j}}}a'_{jk}}$$

3.9. **علاقة المرور** : تسمح لنا علاقة المرور أو العبور من حساب مركبات الأسطرمن مركبة الأسطر نحو الأعمدة أو العكس وتعطى بـ :

$$y_{Jj}^{k}=\frac{1}{\sqrt{\lambda_k}}\sum_{i}\frac{n_{ij}}{n_{\bullet j}}y_{Ii}^{k}$$

$$y_{Ii}^{k}=\frac{1}{\sqrt{\lambda_k}}\sum_{j}\frac{n_{ij}}{n_{i\bullet}}y_{Jj}^{k}$$

مثال 20: بأخذ المثال 19

الولايات / انطباع المستهلك فيما يخص المنتوج	الجزائر	وهران	عنابة	ورقلة	المجموع
ممتاز	9	35	24	8	76
جيد	66	72	112	48	298
متوسط	52	86	43	26	207
ردىء	55	103	40	25	223
المجموع	182	296	219	107	804

تحسب مركبات الأعمدة بأخذ مركبات الأسطر المحسوبة سابقا أي :

الأسطر	y_{Ii}^1	y_{Ii}^2
ممتاز	-0,0419	-0,2926
جيد	0,2972	0,0216
متوسط	-0,1481	0,0510
رديء	-0,2453	0,0235

تحسب مركبات الأعمدة بالنسبة للمحور الأول كمايلي :

λ1 القيمة الذاتية الأولى 0,0552

$$y_{Jj}^k = \frac{1}{\sqrt{\lambda_k}} \sum_i \frac{n_{ij}}{n_{\bullet j}} y_{Ii}^k$$

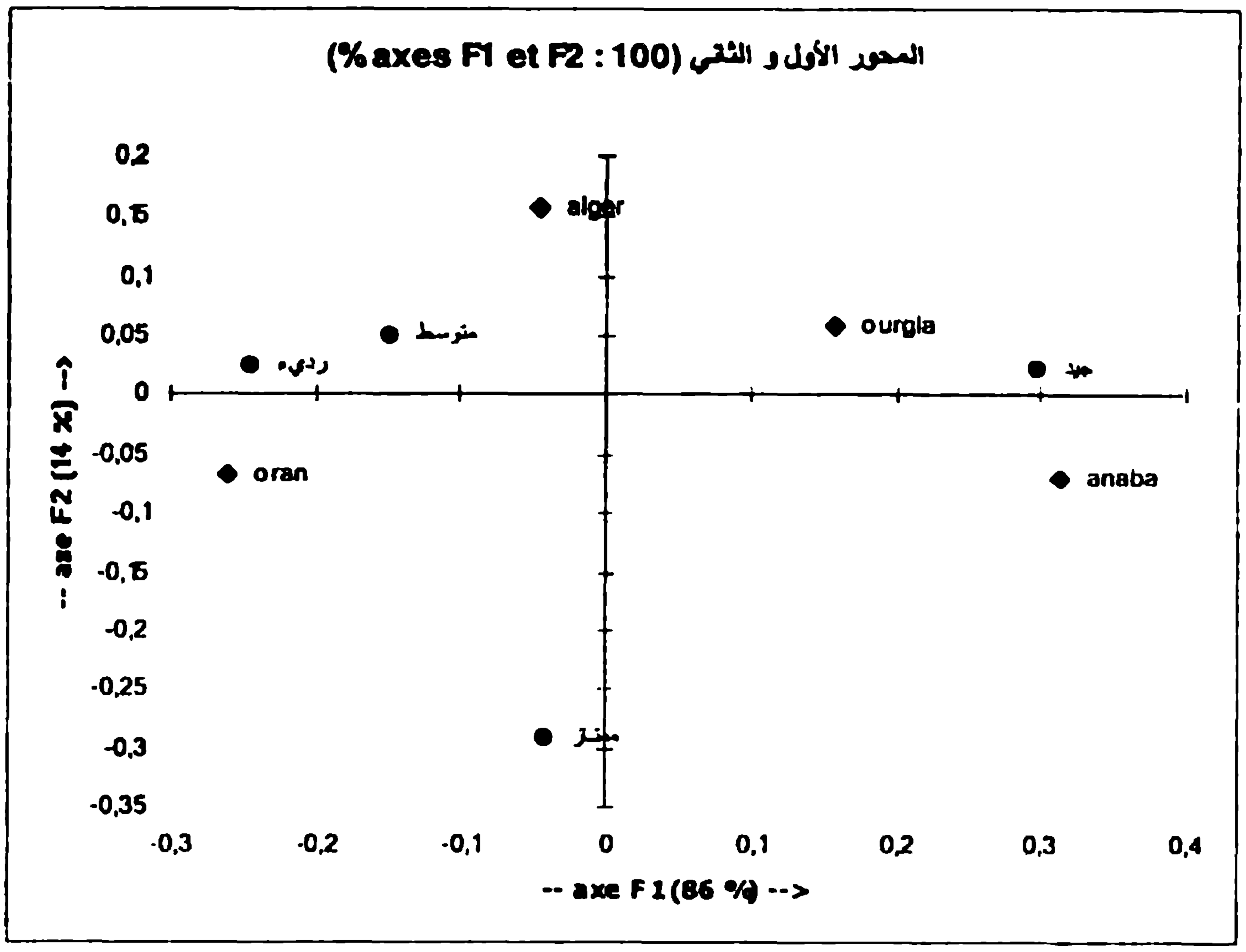
الجزائر	y_{J1}^1	(9/182)*(1/RACINE(0,05524))*(-0,0441)+(66/182)*(1/RACINE(0,05524))*0,294167+ (52/182)*(1/RACINE(0,05524))*-0,149+(55/182)*(1/RACINE(0,05524))*-0,2457	-0,0458

الأعمدة	y_{Jj}^1	y_{Jj}^1
الجزائر	y_{J1}^1	-0,0458
وهران	y_{J2}^1	-0,2599
عنابة	y_{J3}^1	0,3127
ورقلة	y_{J4}^1	0,1568

بنفس الطريقة نحصل على مركبات المحور الثاني للأعمدة:

y_{J1}^2	0,1578
y_{J2}^2	-0,0666
y_{J3}^2	-0,0703
y_{J4}^2	0,0598

تمثيل الأعمدة والأسطر :



من خلال هذا الإسقاط يمكن إعطاء تفسير بدائي لهذا المستوي كمايلي : أغلب الانطباعات الجيدة موجودة في عنابة ، وورقلة ؛ مادام الخاصية " جيد " تقترب منها حسب الشكل، أما الخاصية أو الانطباع " رديء " و" متوسطة " ضعيفة جيدا فيها لأنها

متواجدة في الجهة المعاكس في المحور الأول، ونفس الانطباعات متواجدة بكثرة في وهران، أما انطباع " ممتاز " فهو ضعيف بكثير في الجزائر لأنه موجود في الاتجاه المعاكس في المحور الثاني؛ إلا أنه يجب على هذا التحليل أن يتم عن طريق حساب المساهمات النسبية للأسطر والأعمدة، حيث تسمح هذه المساهمات بتحديد كل محور ومختلف العناصر المحدد له، وبالتالي يصبح التحليل أسهل وأكثر موضوعية ويعطي الصورة الحقيقة للظاهرة.

10. المساهمة المطلقة والمساهمة النسبية

1.10. المساهمة المطلقة والنسبية للأسطر

1.1.10. المساهمة المطلقة للأسطر : إن العلاقة التي تسمح لنا من حساب المساهمة المطلقة لعناصر سحابة الأسطر على المحور k متمثلة في :

$$\alpha_i^K = \frac{n_{i\bullet}}{n_{\bullet\bullet}} \frac{(Y_{li}^k)^2}{\lambda_k}$$

مثال 21 : بأخذ معطيات المثال 18 المتمثلة في :

الأسطر	y_{li}^1	y_{li}^2	ni.	λ1	λ2	n..
ممتاز	-0,0419	-0,2926	76	0,0552	0,0090	804
جيد	0,2972	0,0216	298			
متوسط	-0,1481	0,0510	207			
رديء	-0,2453	0,0235	223			

فان المساهمة المطلقة لعناصر الأسطر في المحور الأول و الثاني بالنسبة المئوية هـي :

$$\alpha_1^1 = \frac{n_{i\bullet}}{n_{\bullet\bullet}} \frac{(Y_{li}^k)^2}{\lambda_k} = \frac{76}{804} \times \frac{(-0.00419)^2}{0.0552} * 100 = 0.3$$

الأسطر	α_i^1	α_i^2
ممتاز	0.3	89.03
جيد	59.29	1.9083
متوسط	10.23	7.3734
ردىء	33.22	1.6785

نلاحظ من خلال المساهمة المطلقة أن الانطباع " الجيد " هو المحدد للمحور الأول، أما الانطباع الممتاز هو المحدد للمحور الثاني.

2.1.10. **المساهمة النسبية للأسطر :** العلاقة التي تسمح لنا من حساب المساهمة النسبية للأسطر على المحور k معطاة بـ :

$$\rho_i^K = \frac{(Y_{li}^k)^2}{\sum_{j=1}^{p} \frac{n_{\bullet\bullet}}{n_{\bullet j}} (\frac{n_{ij}}{n_{i\bullet}} - \frac{n_{\bullet j}}{n_{\bullet\bullet}})^2}$$

باستعمال المركبات الرئيسية للأسطر :

الأسطر	y^1_{li}	y^2_{li}
ممتاز	-0,0419	-0,2926
جيد	0,2972	0,0216
متوسط	-0,1481	0,0510
رديء	-0,2453	0,0235

يتم حساب المساهمة النسبية للممتاز كمايلي :

الأسطر	$\rho^1_i = \dfrac{(Y^k_{li})^2}{\sum_{j=1}^{p} \frac{n_{\bullet\bullet}}{n_{\bullet j}} (\frac{n_{ij}}{n_{i\bullet}} - \frac{n_{\bullet j}}{n_{\bullet\bullet}})^2}$
ممتاز	((-0.0419)*(-0.0419))/(((9/76)-(182/804))^2*(804/182)) + ((35/76)-(296/804))^2*(804/296)) ((24/76)-(219/804))^2*(804/219)) + ((8/76)-(107/804))^2*(804/107)) = 0.0201

بنفس الطريقة نستخرج المساهمة النسبية للأسطر الأخرى :

الأسطر	$\rho^1_i = \dfrac{(Y^k_{li})^2}{\sum_{j=1}^{p} \frac{n_{\bullet\bullet}}{n_{\bullet j}} (\frac{n_{ij}}{n_{i\bullet}} - \frac{n_{\bullet j}}{n_{\bullet\bullet}})^2}$	$\rho^2_i = \dfrac{(Y^k_{li})^2}{\sum_{j=1}^{p} \frac{n_{\bullet\bullet}}{n_{\bullet j}} (\frac{n_{ij}}{n_{i\bullet}} - \frac{n_{\bullet j}}{n_{\bullet\bullet}})^2}$
ممتاز	0.0201	0.9799
جيد	0.9947	0.0053
متوسط	0,8932	0,1059
رديء	0,9907	0,0091

نستنتج من خلال المساهمة النسبية للأسطر أن المحور الثاني هو محور المتازين، حيث أنه يساهم في تحديد المحور الثاني بـ 97,99%، أما المحور الأول هو محور العناصر ذات الخاصية " جيد "، " متوسط "، ورديء

2.10. المساهمة المطلقة و النسبية للعناصر أو الأفراد الموجودة في الأعمدة:

1.2.10. المساهمة المطلقة للأعمدة : العلاقة التي تسمح لنا من حساب المساهمة المطلقة لعناصر سحابة الأعمدة على المحور k هي:

$$\alpha_j^K = \frac{n_{\bullet j}}{n_{\bullet\bullet}} \frac{(Y_{Jj}^k)^2}{\lambda_k}$$

مثال 22 : بأخذ معطيات المثال 21 المتمثلة في :

الأعمدة	y_{Jj}^1	y_{Jj}^2	n.j	λ1	λ2	n..
الجزائر	-0,0458	0,1578	182	0,0552	0,0090	804
وهران	-0,2599	-0,0666	296			
عنابة	0,3127	-0,0703	219			
ورقلة	0,1568	0,0598	107			

إذا المساهمة المطلقة لعناصر الأعمدة في المحور الأول والثاني بالنسبة المئوية هي :

$$\alpha_1^1 = \frac{n_{\bullet j}}{n_{\bullet\bullet}} \frac{(Y_{I_1}^k)^2}{\lambda_k} = \frac{182}{804} \times \frac{(-0.0458)^2}{0.0552} * 100 = 0.86$$

الأعمدة	α_j^1	α_j^2
الجزائر	0.86	61.97
وهران	45.00	17.97
عنابة	48.20	14.80
ورقلة	5.92	5.24

من خلال المساهمة المطلقة للأعمدة فإن الجزائر هي التي تساهم في المحور الثاني، أما وهران و عنابة تساهمان في بناء المحور الأول.

2.2.10. المساهمة النسبية للأعمدة: العلاقة التي تسمح لنا من حساب المساهمة النسبية لعناصر سحابة الأعمدة على المحور k هي :

$$\rho_j^K = \frac{(Y_{Jj}^k)^2}{\sum_{i=1}^{n} \frac{n_{\bullet\bullet}}{n_{i\bullet}} (\frac{n_{ij}}{n_{\bullet j}} - \frac{n_{i\bullet}}{n_{\bullet\bullet}})^2}$$

الأعمدة	$\rho_j^1 = \frac{(Y_{Jj}^k)^2}{\sum_{i=1}^{n} \frac{n_{\bullet\bullet}}{n_{i\bullet}} (\frac{n_{ij}}{n_{\bullet j}} - \frac{n_{i\bullet}}{n_{\bullet\bullet}})^2}$
الجزائر	((-0,0458)*-(0,0458))/((((9/182)-(76/804))^2*(804/76)+ ((66/182)-(298/804))^2*(804/298) ((52/182)-(207/804))^2*(804/207)+ ((55/182)-(223/804))^2*(804/223)))=0,0778

الأعمدة	$\rho_j^1 = \frac{(Y_{Jj}^k)^2}{\sum_{i=1}^{n}\frac{n_{\bullet\bullet}}{n_{i\bullet}}(\frac{n_{ij}}{n_{\bullet j}}-\frac{n_{i\bullet}}{n_{\bullet\bullet}})^2}$	$\rho_j^2 = \frac{(Y_{Jj}^k)^2}{\sum_{i=1}^{n}\frac{n_{\bullet\bullet}}{n_{i\bullet}}(\frac{n_{ij}}{n_{\bullet j}}-\frac{n_{i\bullet}}{n_{\bullet\bullet}})^2}$
الجزائر	0,0778	0,9220
وهران	0,9383	0,0617
عنابة	0,9519	0,0481
ورقلة	0,8713	0,1268

إن الجزائر إسقاطها جيد في المحور الثاني، أما وهران، عنابة، وورقلة إسقاطها جيد في المحور الأول، وبالتالي فإن التحليل الأول الذي قمنا به صحيح، ولكن في الواقع أو في غلب الأحيان نلجأ إلى التحليل بأخذ المحور الثالث لأن نسبة التمثيل لا تأخذ كل العناصر في المحور الأول والثاني ؛ كما هو الحال في مثالنا كون الأسطر والأعمدة إسقاطها جيد في المحور الأول والثاني.

11. النقطة إضافية

في بعض الأحيان يكون من المهم أو من الضروري تمثيل نقاط أو الأعمدة التي لم نقوم بإدخالها في التحليل، لذلك يكفي حساب الإسقاط باستعمال المقياس المختار تحت الفضاء الجزئي.

1.11. إضافة نقطة في العمود : لتكن النقطة الإضافية في العمود التالية :

$$X.J_c = \begin{pmatrix} \frac{n_{1c}}{n_{\bullet c}} \\ \vdots \\ \frac{n_{ic}}{n_{\bullet c}} \\ \vdots \\ \frac{n_{nc}}{n_{\bullet c}} \end{pmatrix}$$

من أجل تمثيل هذا العمود، يكفي حساب إسقاط هذه النقطة باستعمال المقياس المختار تحت الفضاء الجزئي، وعليه فإن المركبة الأساسية k للنقطة الإضافية c يتم الحصول عليها باستعمال علاقة العبور التالية :

$$Y_{Jc}^{K} = \frac{}{\sqrt{\lambda_k}} \sum_{i=1}^{n} \frac{n_{ic}}{n_{\bullet c}} Y_{Ii}^{k}$$

مثال 23 : قمنا بطرح نفس الأسئلة السابقة في ولاية البليدة فكانت النتائج كمايلي :

البليدة	الأسطر
25	ممتاز
19	جيد
38	متوسط
23	رديء
105	المجموع

- أوجد مركبات هذه الولاية ومثلها.

الحل : لدينا مركبات الأسطر من الحساب السابق كما يلي :

الأسطر	y^{1}_{li}	y^{2}_{li}
ممتاز	-0,0419	-0,2926
جيد	0,2972	0,0216
متوسط	-0,1481	0,0510
رديء	-0,2453	0,0235

λ1	λ2
0,0552	0,0090

تكتب النقطة الإضافية للعمود كما يلي :

الأسطر	البليدة
ممتاز	25/105
جيد	19/105
متوسط	38/105
رديء	23/105

$$Y_{Jc}^{K} = \frac{}{\sqrt{\lambda_k}} \sum_{i=1}^{n} \frac{n_{ic}}{n_{\bullet c}} Y_{li}^{k}$$

العمود	$Y_{Jc}^{1}=\frac{}{\sqrt{\lambda_k}}\sum_{i=1}^{n}\frac{n_{ic}}{n_{\bullet c}}Y_{li}^{k}$	$Y_{Jc}^{2}=\frac{}{\sqrt{\lambda_k}}\sum_{i=1}^{n}\frac{n_{ic}}{n_{\bullet c}}Y_{li}^{k}$
البليدة	((25/105)*(-0.0419)+(19/105)*(0.2972)+(38/105)*(-0.1481)+(23/105)*(-0.2453))/(racine(0.0552))=-0,27	((25/105)*(-0.2926)+(19/105)*(0.0216)+(38/105)*(0.0510)+(23/105)*(-0.0235))/(racine(0.0090))=-0,44

ويتم تمثيل هذه النقطة في المستوي السابق و القيام بمعرفة مكانة هذه الأخيرة.

2.11. إضافة نقطة سطريه : لتكن النقطة السطرية الإضافية التاليـــة:

$$XI_s=\begin{pmatrix}\frac{n_{s1}}{n_{s\bullet}}\\ \vdots \\ \frac{n_{sj}}{n_{s\bullet}}\\ \vdots \\ \frac{n_{sp}}{n_{s\bullet}}\end{pmatrix}$$

يكفي حساب إسقاط هذه النقطة ، باستعمال المقياس المختار تحت الفضاء الجزئي من أجل تمثيل هذا السطر ، وعليه فإن المركبة الأساسية k للنقطة الإضافية s يتم الحصول عليها باستعمال علاقة العبور التالية:

$$Y_{Is}^{K}=\frac{}{\sqrt{\lambda_k}}\sum_{j=1}^{p}\frac{n_{sj}}{n_{s\bullet}}Y_{Ij}^{k}$$

ملاحظة : قيمة المساهمة المطلقة تساوي الصفر، بالنسبة لكل من للنقطة الإضافية في السطر أو العمود، بينما المساهمة النسبية فيمكن حسابها، حيث هذه الأخيرة تمثل نسبة تمثيل كل نقطة في محور.

تمارين :

1.2. قنا بسبر للأراء لـ 264 فرد، حيث قمنا بطرح سؤالين Q1، Q2 ؛ حيث كل فرد يقوم باختيار خاصية واحدة من بين أربع خاصيات لكل سؤال، فتحصلنا على الجدول التالي :

السؤال الأول السؤال الثاني	Q1M1	Q1M2	Q1M3	Q1M4	المجموع
Q2M1	20	15	9	7	51
Q2M2	8	11	30	25	74
Q2M3	18	16	12	10	56
Q2M4	6	50	14	13	83
المجموع	52	92	65	55	264

1. هل يمكن القول أنه يوجد ارتباط بين السؤال الأول والسؤال الثاني.

2. أحسب مؤشر التجاذب والتنافر لهذا الجدول. علق النتائج المحصل عليها.

3. أحسب المصفوفة C، حيث المصفوفة $C={}^{t}HH$

$$H = n_{ij} / \sqrt{n_{i\bullet} n_{\bullet j}}$$

4. قمنا بحساب القيم الذاتية للمصفوفة C فتحصلنا على ما يلي :

القيم الذاتية
0,1464
0,1113

أ- استنتج القيم الذاتية الباقية.

ب- أوجد نسبة التمثيل لأحسن مستوي.

5. الشعاع الذاتي المقابل لكل قيمة ذاتية هو كالآتي :

القيم الذاتية	0,1464

الشعاع الذاتي
0,1467
-0,7897
0,4599
0,3788

القيم الذاتية	0,1114

الشعاع الذاتي
-0,8835
0,1644
0,3033
0,3167

أ- أحسب المركبات الأسطر ؛

ب- أحسب المركبات الأعمدة.

6- مثل الأعمدة والأسطر في مستوي.

2.2. أحسب المساهمة النسبية والمطلقة للأسطر وللأعمدة للتمرين الأول.

3.2. قمنا بطرح أربع أسئلة متعلقة بـ 10 مواضيع (إذا كانت لإجابة بـ نعم نضع 1 ، وإذا كانت لا نضع 0)، فكانت النتائج كما يلي :

A4	A3	A2	A1	الأسئلة / الموضوع
1	1	0	0	S1
0	0	1	1	S2
0	1	1	1	S3
1	0	1	0	S4
1	1	0	1	S5
1	0	1	0	S6
0	1	1	1	S7
1	0	1	0	S8
0	1	1	0	S9
1	1	1	1	S10

1. أحسب المصفوفة C، التي تسمح من الحصول على محاور المعاملات، حيث : حيث المصفوفة $C={}^{t}HH$،

و $H = n_{ij} / \sqrt{n_{i\bullet} n_{\bullet j}}$

2. لدينا الأشعة الذاتية التاليـــة :

$$a_1 = \begin{pmatrix} -0.578 \\ 0.121 \\ -0.344 \\ 0.730 \end{pmatrix} ، a_2 = \begin{pmatrix} -0.184 \\ -0.670 \\ 0.663 \\ 0.278 \end{pmatrix} ، a_3 = \begin{pmatrix} 0.658 \\ -0.465 \\ -0.450 \\ 0.386 \end{pmatrix}$$

القيم الذاتية المرافق لهذه الأشعة هـي :

$$\lambda_1 = 0.303 ، \lambda_2 = 0.181 ، \lambda_3 = 0.097$$

- استنتج القيمة الذاتية الباقية. ثم أحسب نسبة التمثيل للمستوي.

- مثل الأسئلة والمواضيع في مستوي.

3. قمنا بطرح سؤال جديد فكانت النتائج كمايلـي :

A5	1	1	0	1	1	1	0	1	0	0

4. أوجد مركبات هذا السؤال

5. قمنـا بإضـافة موضـع آخـر علـى الجـدول الـسابق فكانـت النتائج كمايلي :

1	0	1	0	S11

أوجد مركبات هذا الموضوع.

الفصل الثالث

التصنيف التسلسلي

Classification hiérarchique

1. تمهيد

بعد القيام بإسقاط الأفراد والمتغيرات في المستوي عند استعمال طريقة التحليل المعطيات بالمركبات الرئيسية، أو عند القيام تطبيق طريقة التحليل بالمعاملات للجدول مزدوج ؛ غالبا ما ندعم تحليلنا عن طريق تصنيف الأفراد أو المتغيرات، حيث يتمثل التصنيف التسلسلي في إعطاء لمجموعة " E " مجموعة من الأقسام، تم الحصول عليها عن طريق التجميع المتتالي لأجزاء أو عناصر كانت متشابهة أم لا ؛ غير أنه يوجد تقنيتان للتصنيف :

- التصنيف غير المتسلسل أو ما يسمى بالتجزئة: يتمثل هذا الأخير في تجزئة المجموعة الكلية للأفراد إلى n مجموعة منفصلة ؛ عدد المجموعات محدد بـ n.

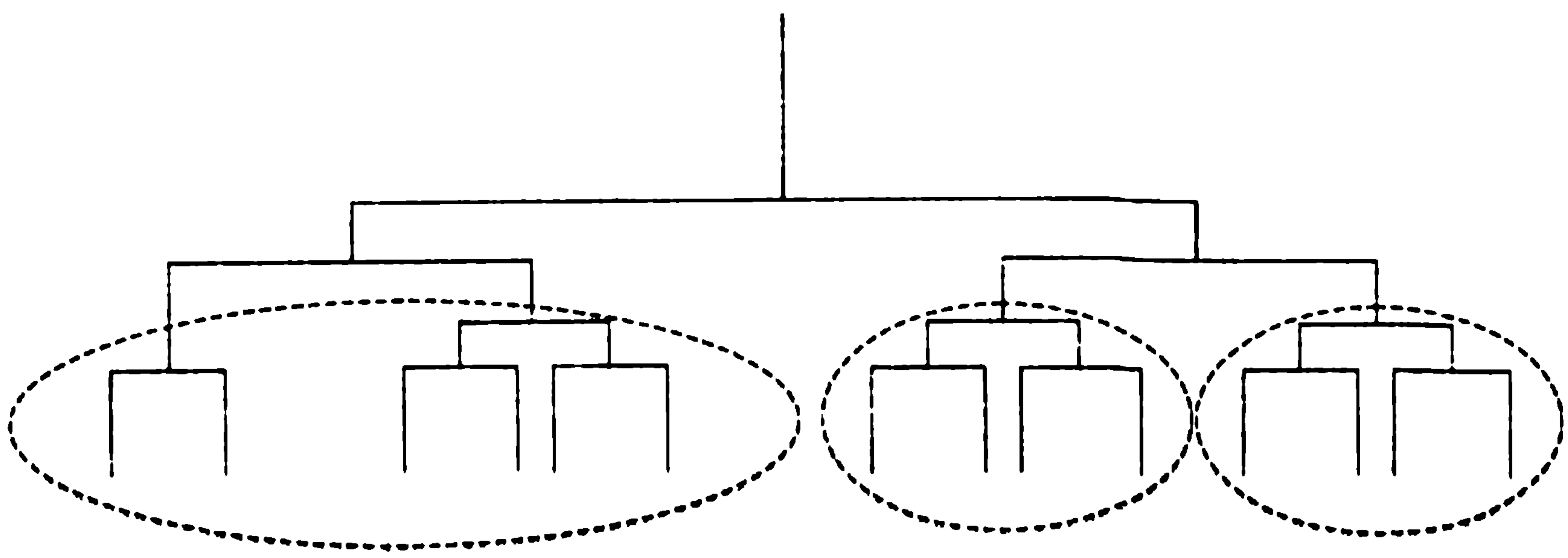
- التصنيف التسلسلي : المتمثلة في تجميع الأفراد إلى مجموعات.

2. مبدأ وهدف التصنيف التسلسلي

يعتمد مبدأ التصنيف التسلسلي في إنشاء في كل مرحلة من التجزئة على جمع العناصر مثنى مثنى، للعناصر الأكثر اقترابا، ويتم التصنيف عن طريق شجرة تسلسلية تسمى بـ دندوقرام " Dendrogramme "، هذه الشجرة عادة ما يتم الحصول عليها في أغلب الطرق المستعملة عن طريق طريقة تصاعدية ؛ حيث نقوم بتجميع أولاً الفردين الأكثر اقترابا والتي تشكل قمة، وبالتالي يتبقى n-1 هدف

(Objets) ثم نقوم بتكرار العملية حتى نقوم بتجميع الكلي، ومن المشاكل التي تواجه هذه العملية لدينا تعريف مقياس عدم التشابه بين المجموعات.

الشكل 3 : شجرة دندوقرام Dendrogramme



نقوم بتصنيف الأفراد من أجل تجميع على سبيل المثال :

أ- زبائن مؤسسة: لتكن مجموعة من الزبائن لمؤسسة موصوفة بنوعية المشتريات، والمتغيرات السوسيولوجيا ؛ إذا حسب المعلومة التي نريد الحصول عليها، يمكن أن نقوم بتصنيف الزبائن حسب نوعية المشتريات أو حسب المتغيرات السوسيولوجيا.

ب- مجموعة من الدول : لتكن مجموعة من الدول موصوفة بمجموعة من المتغيرات الاقتصادية والديمغرافيا ؛ وعليه يمكن تصنيف الدول حسب التحليل أو النتائج المراد الحصول عليها، إما حسب المؤشرات الاقتصادية أو الديمغرافيا أو الاثنين معا.

أما بالنسبة للمتغيرات يمكن تصنيفها من أجل تقليص عددها، حيث يمكن أن يكون عدد المتغيرات معتبر ومتكررة أي هناك متغيرات تفسر بنفس القدر ؛ وبالتالي يسمح التصنيف من اختيار المتغيرات الأكثر تمثيلا، وعن طريق هذه المجموعة الجديدة من المتغيرات يمكن تصنيف الأفراد.

3. مقياس التفرقة بين الأفراد

نستعمل المسافة مابين الأفراد من أحل التفرقة بين الأفراد، ولقياس هذه المسافة يمكن استعمال المؤشرات التالية :

1.3. **المسافة الاقليدية :** يعتبر هذه النوع من المسافات الأكثر استعملا، وهي متمثلة في مسافة هندسية لفضاء متعدد الأبعاد وهي تعطى بالعلاقة التالية :

$$d(i,i')=\sqrt{\sum_{j=1}^{p}\left(x_{ij}-x_{i'j}\right)^2}$$

2.3. **المسافة الاقليدية مربعة :** وهي متمثلة في تربيع المسافة الاقليدية، حيث أنها تسمح من أخذ بعين الاعتبار الأفراد الشاذة (أي المبتعدة كثيرا) و هي معطاة بـ :

$$d^2(i,i')=\sum_{j=1}^{p}\left(x_{ij}-x_{i'j}\right)^2$$

مثال 24 : لدينا الجدول التالي :

	X1	X2	X3	X4
A	5	10	13	4
B	16	7	6	1
C	8	15	14	3
D	9	2	11	12

أحسب جدول المسافات بين الأفراد ، باستعمال المسافة الاقليدية مربعة.

الحل : يتم حساب المسافة مابين الأفراد بالعلاقة التالية :

$$d^2(i,i')=\sum_{j=1}^{p}\left(x_{ij}-x_{i'j}\right)^2$$

جدول المسافات للعناصر الأساسية :

	A	B	C	D
A	0	188	36	148
B	188	0	196	220
C	36	196	0	260
D	148	220	260	0

$$d^2(A,B)=(5-16)^2+(10-7)^2+(13-6)^2+(4-1)^2=188$$

4. مراحل التصنيف

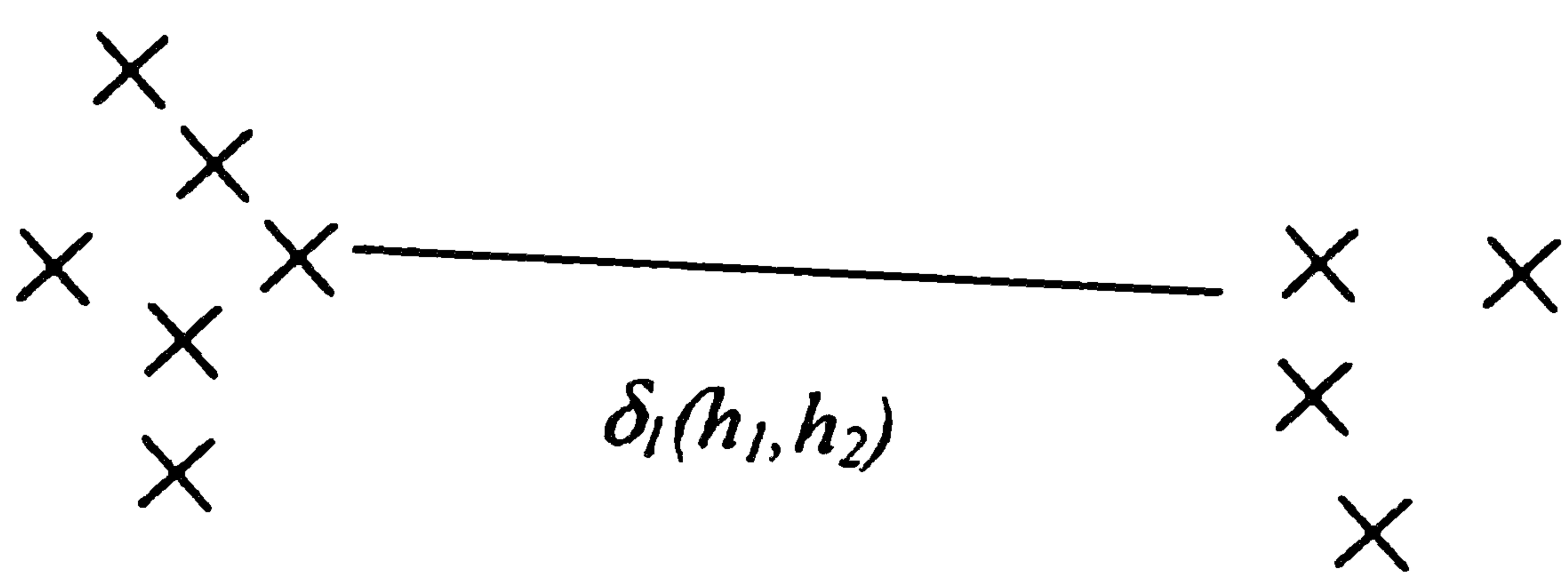
تتمثل مراحل التصنيف في أخذ كل عناصر المجموعة الكلية ، وبحساب المسافات فيما بينها ، ثم نقوم باختيار مؤشر التجميع ؛ بعد

اختيار مؤشر التجميع δ، يتم اختيار الأقسام والأصناف التي تشكل التصنيف H المراد الحصول عليه كمايلي :

1. التصنيف حسب العناصر؛
2. بناء تصنيف جديد عن طريق تجميع صنفين أو قسمين من التصنيف السابق التي تصغر δ ؛
3. إعادة العملية 2 حتى يتم تجميع كل الأصناف في مجموعة واحدة.

ومن مؤشرات التجميع لدينا :

1.4. **مؤشر التجميع لأدنى مسافة :** يعتبر مؤشر التجميع لأدنى مسافة من أبسط مؤشرات التجميع، وهي متمثلة في أقرب زوج من العناصر كما هو موضح في الشكل الموالي :



$$\delta(h_1, h_2) = \underset{\substack{x_i \in h_1 \\ x_j \in h_2}}{Min}\, d^2(x_i, x_j)$$

حيث نقوم بضم فرد x والمجموعة h_2 كمايلي :

$$\delta(\{x_1\}, h_2) = \underset{x \in h_2}{Min}\, d^2(x_1, x)$$

ومن عيوب هذا المؤشر أن يقوم بالتجميع العناصر الواحد تلوى الآخر، ولكن من ايجابيته أي يشوه قليلا المجموعة الأصلية.

مثال 25 : بأخذ نفس مصفوفة المسافات للمثال 24.

جدول المسافات للعناصر الأساسية.

	A	B	C	D
A	0	188	**36**	148
B	188	0	196	220
C	36	196	0	260
D	148	220	260	0

أدنى مسافة في هذه المصفوفة هي 36 إذا نقوم بتجميع العنصر C مع A في مجموعة نسميها E .

يتم الحصول على المسافة بين المجموعة E والعناصر الأخرى كمايلي :

$$\delta(E,B) = Min\left[\mathrm{d}^2(\mathrm{A,B}); d^2(C,B)\right] = d^2(A,B) = 188$$
$$\delta(E,D) = Min\left[\mathrm{d}^2(\mathrm{A,D}); d^2(C,D)\right] = d^2(A,D) = 148$$

نحصل على المصفوفة التالية :

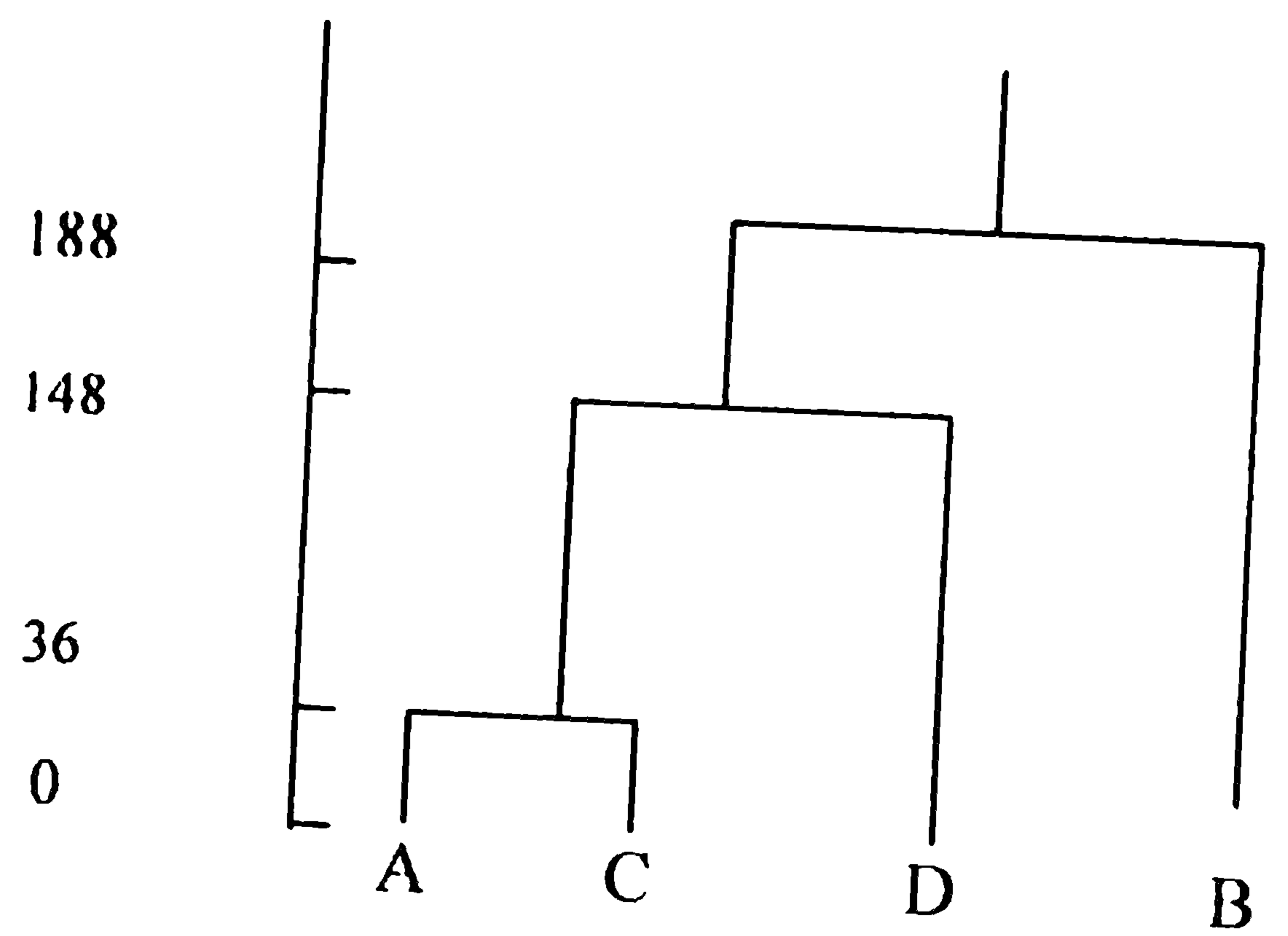
	B	D	E
B	0	220	188
D	220	0	148
E	188	148	0

نقوم من خلال هذه المصفوفة بتجميع الفرد D مع المجموعة E فنحصل على المجموعة F، وبتالي يتم الحصول على مصفوفة المسافات كمايلي ":

$$\delta(F,B) = Min\left[d^2(E,B); d^2(D,B\right] = d^2(D,B) = 188$$

	B	F
B	0	188
F	188	0

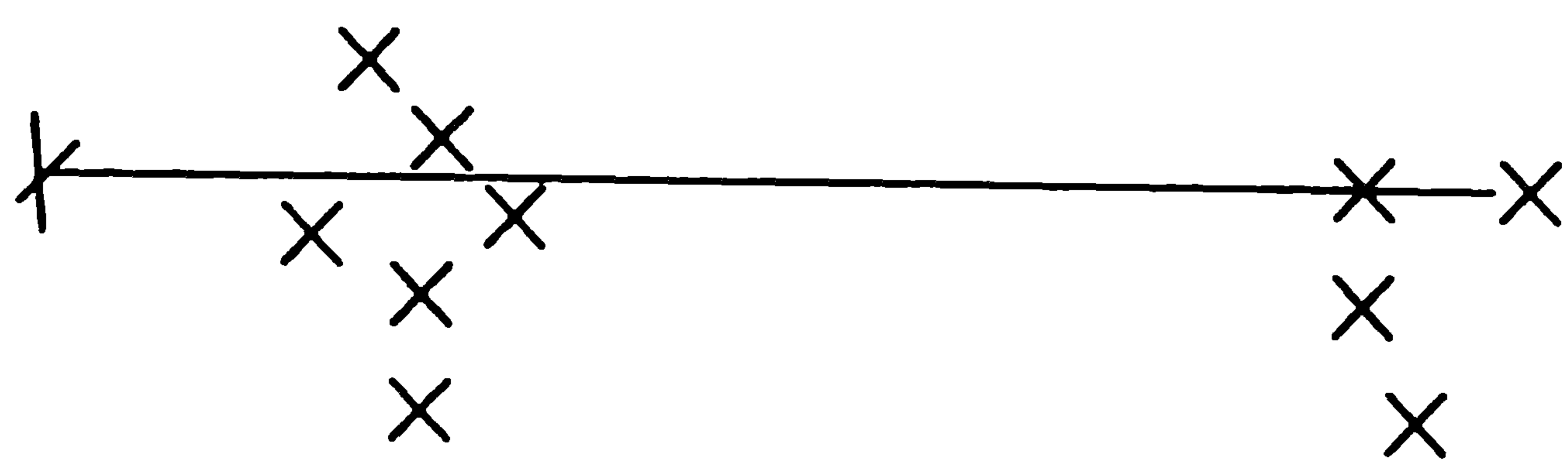
ومنه شكل الدندوقرام هو :



من خلال هذا الدندوقرام نقوم بتجميع (A, C) وهي تكون المجموعة الأولى وD مع المجموعة السابقة تشكل المجموعة الثانية، ومع إضافة العنصر B نشكل المجموعة الكلية أي كما قمنا بذكره سابقا فإن هذا المؤشر يقوم بجمع عنصر بعنصر.

يمكن الحصول على شجرة الدندوقرام عن طريق برنامج الآلي SPSS انظر الملحق 4 للمراحل تطبيق التصنيف على الحسوب باستعمال SPSS.

2.4. مؤشر التجميع لأقصى مسافة : هذا المؤشر معاكس للمؤشر السابق، حيث يتم الجمع بين المجموعات عن طريق أقصى مسافة بين زوج من الأفراد، وهذا موضح في الشكل الموالي :



$$\delta(h_1, h_2) = \underset{\substack{x_i \in h_1 \\ x_j \in h_2}}{Max}\, d^2(x_i, x_j)$$

حيث نقوم بضم فرد x والمجموعة h_2 كمايلي :

$$\delta(\{x_1\}, h_2) = \underset{x \in h_2}{Max}\, d^2(x_1, x)$$

من عيوب هذه الطريقة أنها تشوه كثيرا المجموعات الأصلية.

مثال 26 : لتكن مصفوفة المسافات التالية :

	A	B	C	D
A	0	188	36	148
B	188	0	196	220
C	36	196	0	260
D	148	220	260	0

أدنى مسافة في هذه المصفوفة هي 36، إذا نقوم بتجميع العنصر C مع A في مجموعة نسميها E.

يتم الحصول على المسافة بين المجموعة E، والعناصر الأخرى كمايلي :

$$\delta(E,B) = Max\left[d^2(A,B); d^2(C,B)\right] = d^2(C,B) = 196$$

$$\delta(E,D) = Max\left[d^2(A,D); d^2(C,D)\right] = d^2(A,D) = 260$$

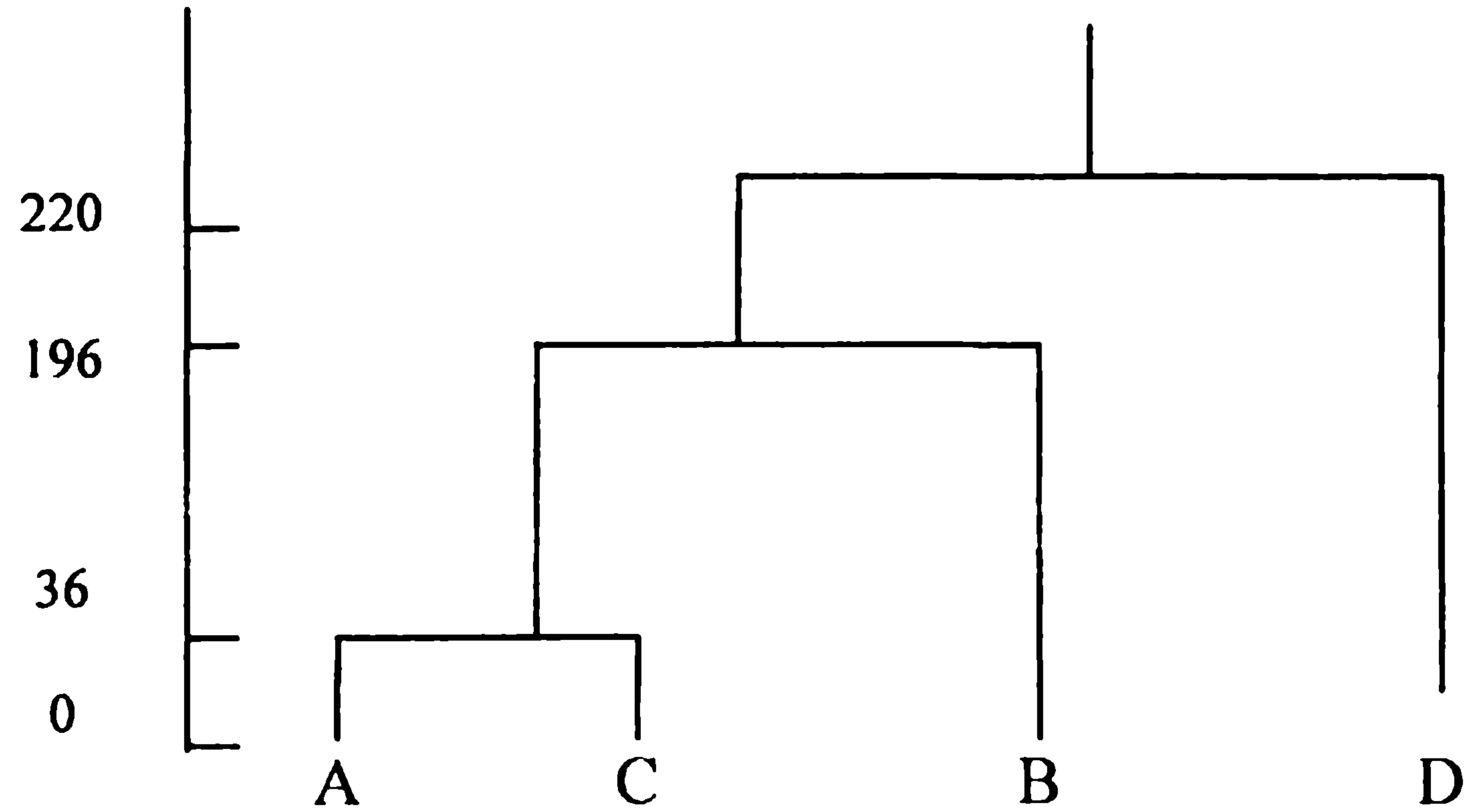
فنحصل على المصفوفة التاليـــة :

	B	D	E
B	0	220	196
D	220	0	260
E	**196**	260	0

نقوم من خلال هذه المصفوفة بتجميع الفرد B مع المجموعة E فنحصل غلى المجموعة F، وبتالي مصفوفة المسافات يتم الحصول عليها كما يلي :

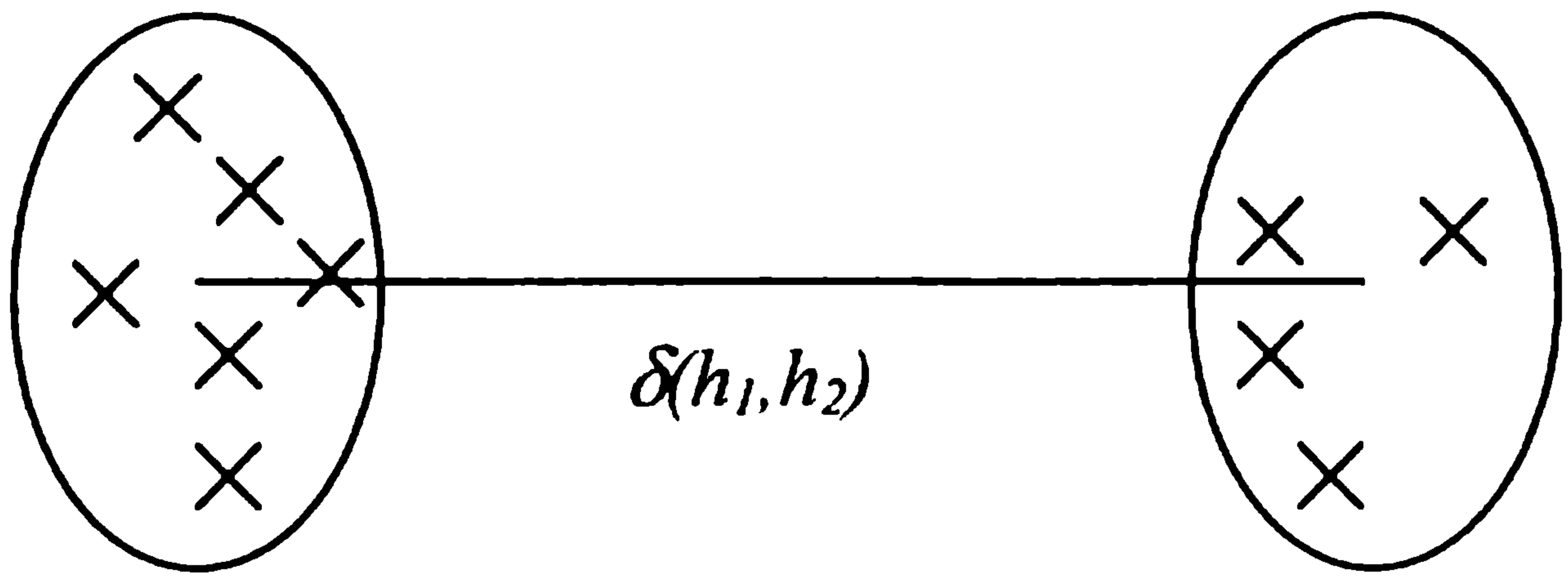
$$\delta(F,D) = Max\left[d^2(E,D); d^2(D,B)\right] = d^2(D,B) = 220$$

	D	F
D	0	220
F	220	0



نقوم من خلال هذا الدندوقرام بتجميع (A, C) وهي تكون المجموعة الأولى وB مع المجموعة السابقة تشكل المجموعة الثانية، ومع إضافة العنصر D نشكل المجموعة الكلية. نلاحظ أن هذا المؤشر يختلف في تجميع العناصر B وD بالمقارنة مع مؤشر التجميع للأدنى مسافة.

3.4. **مؤشر التجميع للمسافة المتوسطة :** يعتمد هذا المؤشر على متوسط الأفراد في المجموعة، أي يتم عن طريق متوسط فرد ومجموعة أو بين متوسط الأفراد في مجموعة مع مجموعة أخرى كما هو موضح في الشكل الموالي :



$$\delta(h_1,h_2)=\underset{\substack{x_i\in h_1\\x_j\in h_2}}{Moy}\, d^2(x_i,x_j)$$

مثال 27 : لتكن مصفوفة المسافات التاليـــة :

	A	B	C	D
A	0	188	**36**	148
B	188	0	196	220
C	36	196	0	260
D	148	220	260	0

أدنى مسافة في هذه المصفوفة هي 36 ؛ إذا نقوم بتجميع العنصر C مع A في مجموعة نسميها E ، فنحصل على الجدول التالي :

	B	D	E
B	0	220	192
D	220	0	204
E	**192**	204	0

$$\delta(h_1,h_2)=\underset{\substack{x_i\in h_1\\x_j\in h_2}}{Moy}\, d^2(x_i,x_j)$$

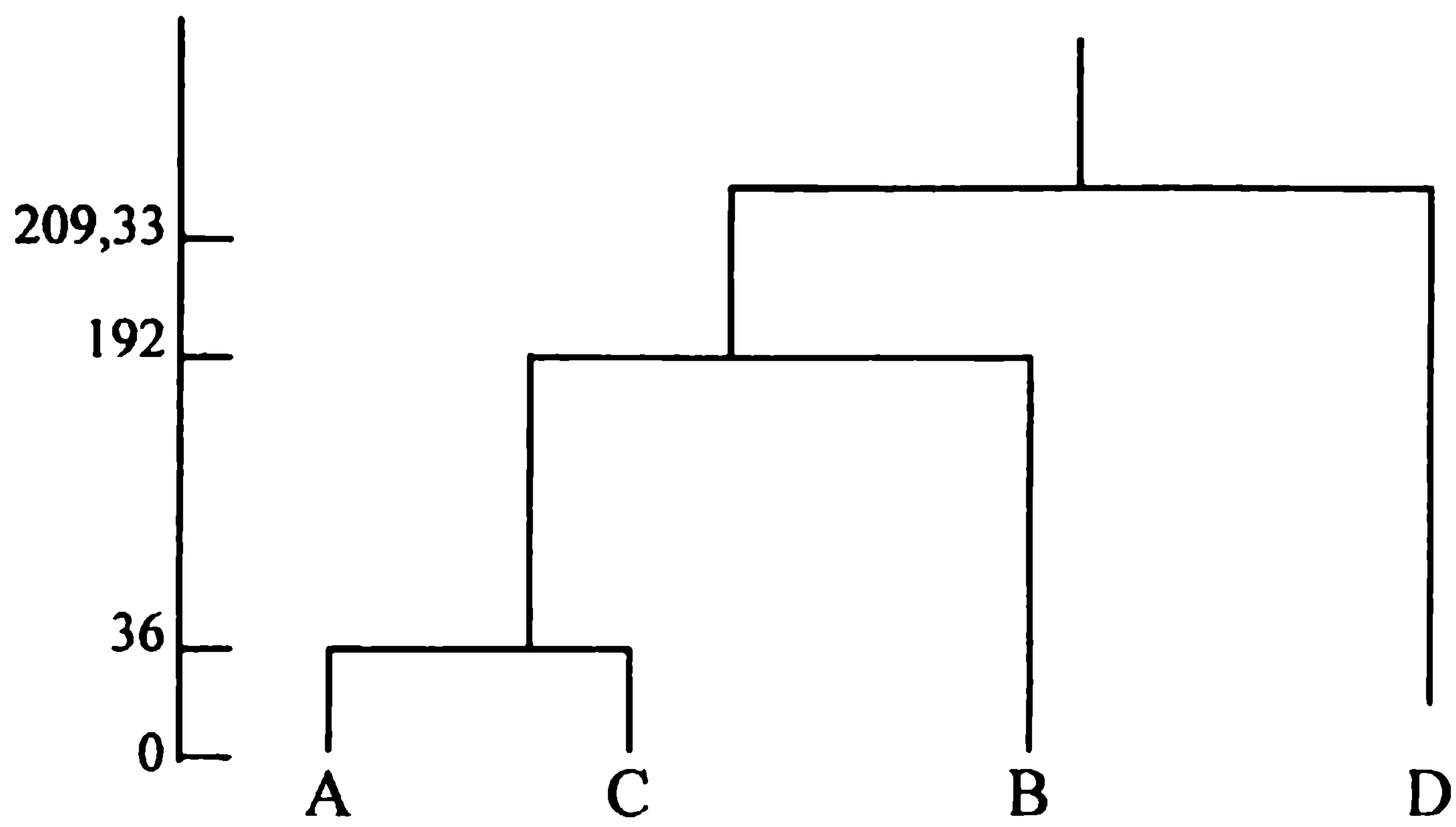
$$\delta_2(E,B)=Moy\{d^2(A,B),d^2(C,B)\}\Leftrightarrow Moy\{(188),(196)\}=192$$

$$\delta_2(E,D)=Moy\{d^2(A,D),d^2(C,D)\}\Leftrightarrow Moy\{(148),(260)\}=204$$

أدنى مسافة في هذه المصفوفة هي مسافة 192 . إذا نقوم بتجميع العنصر B مع المجموعة E في مجموعة نسميها F ، فنحصل على الجدول التالي :

	D	F
D	0	209.33
F	209.33	0

$$\delta_2(F,D)=Moy\{d^2(E,D),d^2(B,D)\}\Leftrightarrow Moy\{(148),(260),(220)\}=209.33$$



نقوم من خلال هذا الدندوقرام بتجميع (A, C) وهي تكون المجموعة الأولى وB مع المجموعة السابقة تشكل المجموعة الثانية ، ومع إضافة العنصر D نشكل المجموعة الكلية. نلاحظ أن هذا المؤشر يختلف في تجميع العناصر B وD بالمقارنة مع مؤشر التجميع للأدنى مسافة.

نظرا لخصوصيات المؤشرات السابقة، المتمثلة في تجميع عنصر بعنصر، وعدم إمكانية الحصول على تفكيك إلى مجموعات منفصلة تماما، سنتطرق إلى مؤشر التجميع يعتمد على مركز الثقل مثل مؤشر WARD.

4.4. مؤشر التجميع لـ WARD : اقترح Ward (1963) طريقة تسمح من الحصول على مجموعات بطريقة تصغر من فقدان المعلومات الناتجة من كل تجميع حيث أنها تعتمد على تحليل التباين، حيث يتمثل في تدنئه أو تصغير مجموع المربعات لكل الأزواج الممكن تشكيلها في كل مرحلة، حيث أن مؤشر الفصل أو التفرقة هو يساوي إلى الكمية المفقودة من التجميع، لنفترض أننا لدينا مركز الثقل للمجموعتين g_A و g_B فإن مركز الثقل لجمع المجموعتين هو :

$$g_{AB} = \frac{n_a g_a + n_b g_b}{n_a + n_b}$$

التباين مابين المجموعات هو متوسط المربع لبعد المسافة لمركز الثقل المجموعات عن مركز الثقل الكلي، يساوي التغير في تباين مابين المجموعات، عند تجميع A و B :

$$n_A d^2(g_A, g) + n_B d^2(g_B, g) - (n_A + n_B) d^2(g_{AB}, g)$$

أي :

$$\delta(A,B) = \frac{n_A n_B}{n_A + n_B} d^2(g_A, g_B)$$

تـدخل هـذه الطريقـة في إطـار علاقـة Lance et Williamas المعممة، جمع المجموعة (A, B) والعنصر C ؛ يتم كمايلي :

$$\delta[(A,B);C]=\frac{(n_A+n_B)\delta(A,B)+(n_B+n_C)\delta(B,C)-n_C\delta(A,B)}{n_A+n_B+n_C}$$

تعتـبر طريقـة Ward ، أحـسن طريقـة تجميـع مـن طـرف بعـض الإحصائيين، حيث أنها تسمح من تحديد مجموعات مجزئة ومنفصلة.

مثال 28 : لتكن مصفوفة المسافات التاليـة :

	A	B	C	D
A	0	188	36	148
B	188	0	196	220
C	36	196	0	260
D	148	220	260	0

يتم تصنيف المجموعات عن طريق طريقة Ward كما يلي :

- أولا : يتم حساب مصفوفة التجميع عن طريق العلاقة التالية :

$$\delta(A,B)=\frac{n_A n_B}{n_A+n_B}d^2(g_A,g_B)$$

نحصل علـى :

	A	B	C	D
A	0	94	**18**	74
B	94	0	98	110
C	18	196	0	130
D	74	110	130	0

أدنى مسافة 18. إذا نقوم بتجميع العنصر C مع A في مجموعة نسميها E ، ونقوم بحساب التغير في التباين باستعمال العلاقة التالية :

$$\delta[E(A,C);D]=\frac{(n_A+n_D)\delta(A,D)+(n_C+n_D)\delta(C,D)-n_D\delta(A,C)}{n_A+n_C+n_D}$$

$$=\frac{(1+1)(74)+(1+1)(130)-(1)(18)}{1+1+1}=130$$

$$\delta[E(A,C);B]=\frac{(n_A+n_B)\delta(A,B)+(n_C+n_B)\delta(C,B)-n_D\delta(A,C)}{n_A+n_C+n_B}$$

$$=\frac{(1+1)(94)+(1+1)(98)-(1)(18)}{1+1+1}=122$$

	E	B	D
E	0	122	130
B	122	0	110
D	130	**110**	0

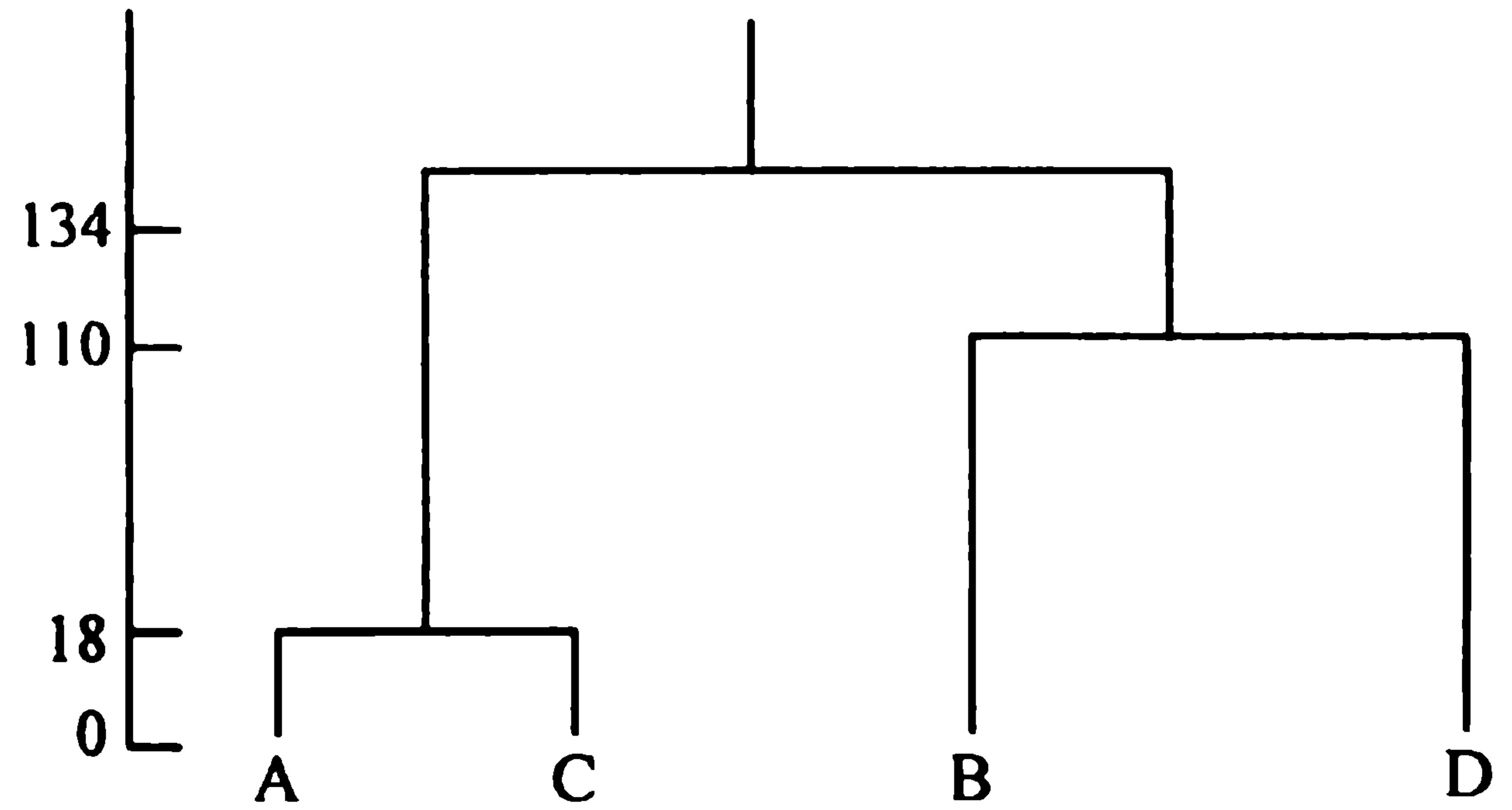
أدنى مسافة 110. إذا نقوم بتجميع العنصر D مع المجموعة E نسميها F ، نقوم بحساب التغير في التباين باستعمال العلاقة التالية :

$$\delta[F(B,D);E]=\frac{(n_B+n_E)\delta(B,E)+(n_D+n_E)\delta(D,E)-n_D\delta(B,D)}{n_F+n_E}$$

$$=\frac{(2+1)(122)+(1+2)(130)-(1)(18)}{2+2}=134$$

نحصل على الجدول التالـي :

	E	F
E	0	134
F	134	0



من خصوصيات التصنيف عن طريق طريقة WARD أنها تسمح من تفادي التصنيف المتسلسل للعناصر ؛ أي أنها تسمح من الحصول على المجموعات منفصلة كما هو ملاحظ في هذا المثال، بالإضافة إلى ذلك عادة ما نستعمل هذه الطريقة على المركبات الرئيسية عند القيام بتطبيق طريقة ACP أو ACF.

تمارين

1.3. لدينا الجدول التالي المتعلق بالكمية المطلوبة من السلع من طرف 6 زبائن :

السلع / الزبائن	X_1	X_2	X_3	X_4	X_5
A	7	8	7	5	17
B	8	9	12	10	10
C	9	6	13	8	9
D	10	7	10	7	11
E	11	10	9	6	12
F	12	11	15	8	14

1. أحسب جدول المسافات مربعة للأفراد.

2. أرسم الدندوقرام للأفراد باستعمال :

أ- مؤشر التجميع لأدنى مسافة.

ب- مؤشر التجميع لمتوسط المسافة.

ج- مؤشر التجميع لـ WARD.

2.3. بأخذ معطيات التمرين السابق :

1- أحسب جدول المسافات مربعة للمتغيرات.

2- صنف المتغيرات باستعمال :

أ- مؤشر التجميع لأدنى مسافة.

ب- مؤشر التجميع لـ WARD.

3.3. لتكن المعطيات المتعلقة بمبيعات مؤسسة لأربع منتوجاتها في الولايات الجزائرية التالية :

المبيعات الولايات	X1	X2	X3	X4
الجزائر	150	130	60	30
ورقلة	60	50	30	20
البليدة	40	30	20	40
بومرداس	30	40	15	16
قسنطينة	90	100	70	17
عنابة	180	70	50	18
وهران	120	50	30	20

قمنا بتطبيق ACP المعيارية على هذه الولايـات 88,37%.

1- أحسب مصفوفة المسافات للمركبات الرئيسيـة.

2- صنف هذه الولايات باستعمال طريـقة Ward.

حـــل التماريـــن

الفصـــل الأول

-1

X3	X2	X1	الأفراد (الولايات)
70	15	30	الجزائر العاصمة
55	20	14	وهران
40	17	25	عنابة
30	14	31	ورقلة
48,75	16,5	25	المتوسط

-2

الجدول المركز

X1	X2	X3
5	-1,5	21,25
-11	3,5	6,25
0	0,5	-8,75
6	-2,5	-18,75

-2

X'

5	-11	0	6
-1,5	3,5	0,5	-2,5
21,25	6,25	-8.75	-18,75

1/4 X/X			
	45,5	-15,25	-18,75
1/4 X/X	-15,25	5,25	8,125
	-18,75	8,125	229,6875

-3

trace (1/4X/X) = 280,4375

$\lambda 2=(trcae1/n(x/x)-(\lambda 1+\lambda 3))$ = 48,34

4- نسبة تمثيل المحاور :

$$cr(\Delta_1 \oplus \Delta_2 / I_G) = \frac{\lambda_1+\lambda_2}{\lambda_1+\lambda_2+...+\lambda_p} = \frac{\lambda_1+\lambda_2}{trace(1/nX'X)}$$

% 99.96

5- حساب المركبات الرئيسية :

يجب التحقق من أن طويلة الأشعة الذاتية تساوي الواحد حتى تشكل أساس :

	-0,1
a1	0,04
	0,99

$$\|a1\|^2 = (-0.10)^2 + (0.04)^2 + (0.99)^2 = 0.9917$$

الطويلة تساوي الواحد

	0,94
a2	-0,31
	0,11

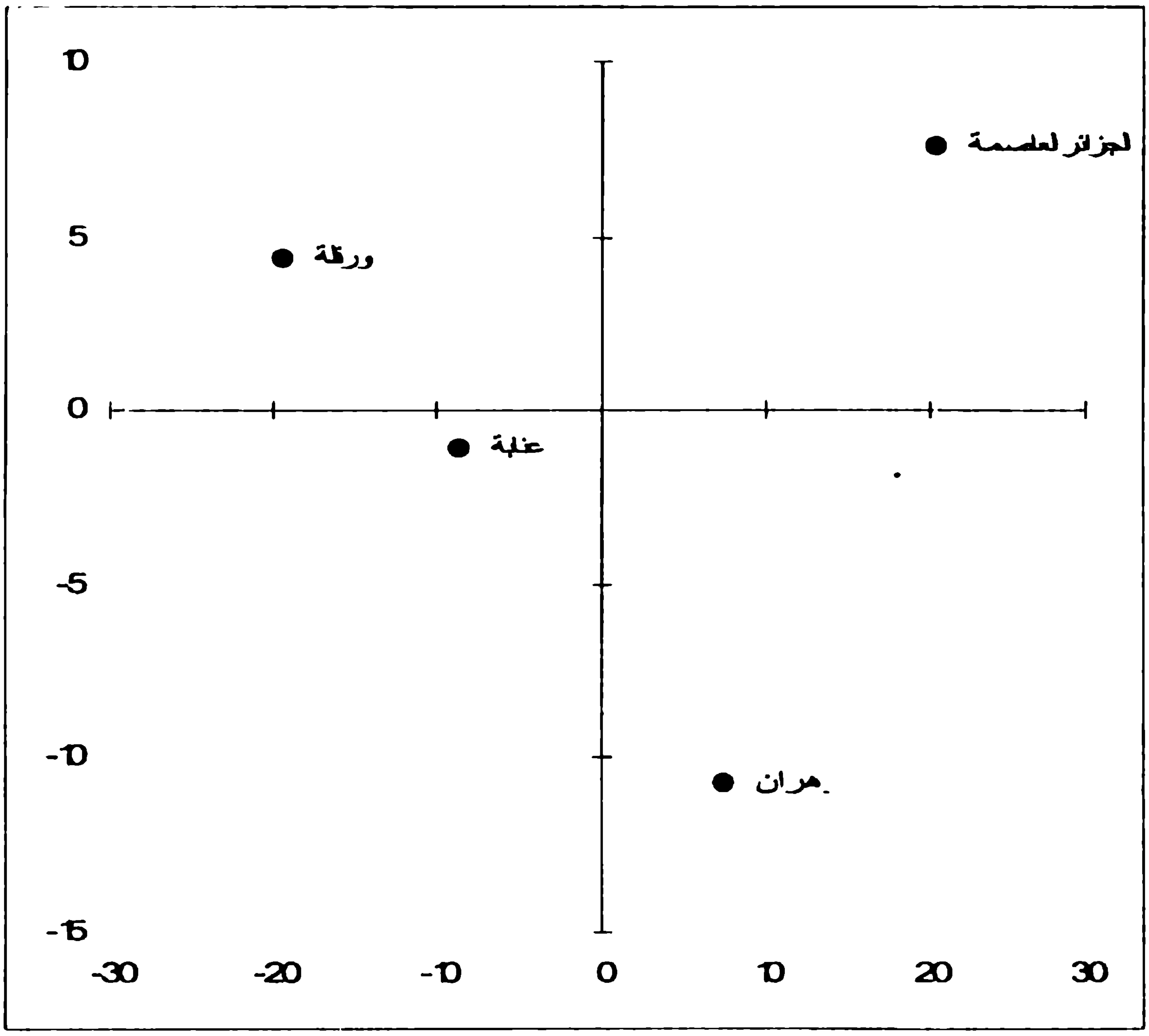
نفس الشيء بالنسبة للمحور الثاني، إذا يتم الاعتماد على هذه الأشعة لحساب مركبات الأفراد.

أ- حساب مركبات الأفراد :

$$Y_i = {}^t A U_{ci}$$

ورقلة	عنابة	وهران	الجزائر	الأفراد
-19,26	-8,64	7,43	20,48	المركبات في المحور 1
4,35	-1,12	-10,74	7,50	المركبات في المحور 2

التمثيل البياني :



.2.1

الأفراد	cos2(α1)	cos2(α2)	المستوي cos2(α)	Cos2(α3)
الجزائر	**0,8757**	**0,1175**	**0,9933**	0,0066
عنابة	**0,3201**	**0,6690**	**0,9892**	0,0107
وهران	**0,9724**	**0,0162**	**0,9886**	0,0113
ورقلة	**0,9421**	**0,0481**	**0,9902**	0,0097

إذا كل من الجزائر، وهران وورقلة ممثلة جيدا في المحور الأول، أما عنابة ممثلة جيدا في المحور الثاني.

2- نحن لدينا

القيم الذاتية	
213,99	الأولى
48,34	الثانية
0,11	الثالثة

المركبات في المحور الثاني	المركبات في المحور الثاني
20,54	7,56
7,5	-10,77
-8,67	-1,13
-19,36	4,35

مساهمة الأفراد هي :

الأفراد	المحور 1	المحور الثاني
الجزائر	0,4928	0,2955
عنابة	0,0657	0,5998
وهران	0,0878	0,0066
ورقلة	0,4378	0,0978

مساهمة الأفراد بالنسبة المؤية % :

الأفراد	المحور 1	المحور الثانى
الجزائر	49,28	29,55
عنابة	6,57	59,98
وهران	8,78	0,66
ورقلة	43,78	9,78

إذا كل من الجزائر وورقلة هي ممثلة للمحور الأول، أما المحور الثاني مشكل من الجزائر وعنابة.

يتم الحصول على المساهمة النسبية في المحور الثالث باستعمال ما يلي :

a3	0,32	0,95	-0,01

X1	5	-11	0	6
X2	-1,5	3,5	0,5	-2,5
X3	21,25	6,25	-8,75	-18,75

الأفراد	الإحداثيات في المحور الثالث
الجزائر	-0,0375
عنابة	-0,2575
وهران	0,5625
ورقلة	-0,2675

إذا مساهمة الأفراد في المحور الثالث هي :

	المحور الثالث	المحور الثالث %
الجزائر	0,0031	**0,31**
عنابة	0,1506	**15,06**
وهران	0,7191	**71,91**
ورقلة	0,1626	**16,26**

حسب هذه المساهمة وهران هي التي تحدد المحور الثالث

-3

D1/σ			
	0,1482	0	0
D1/σ	0	0,4364	0
	0	0	0,0659

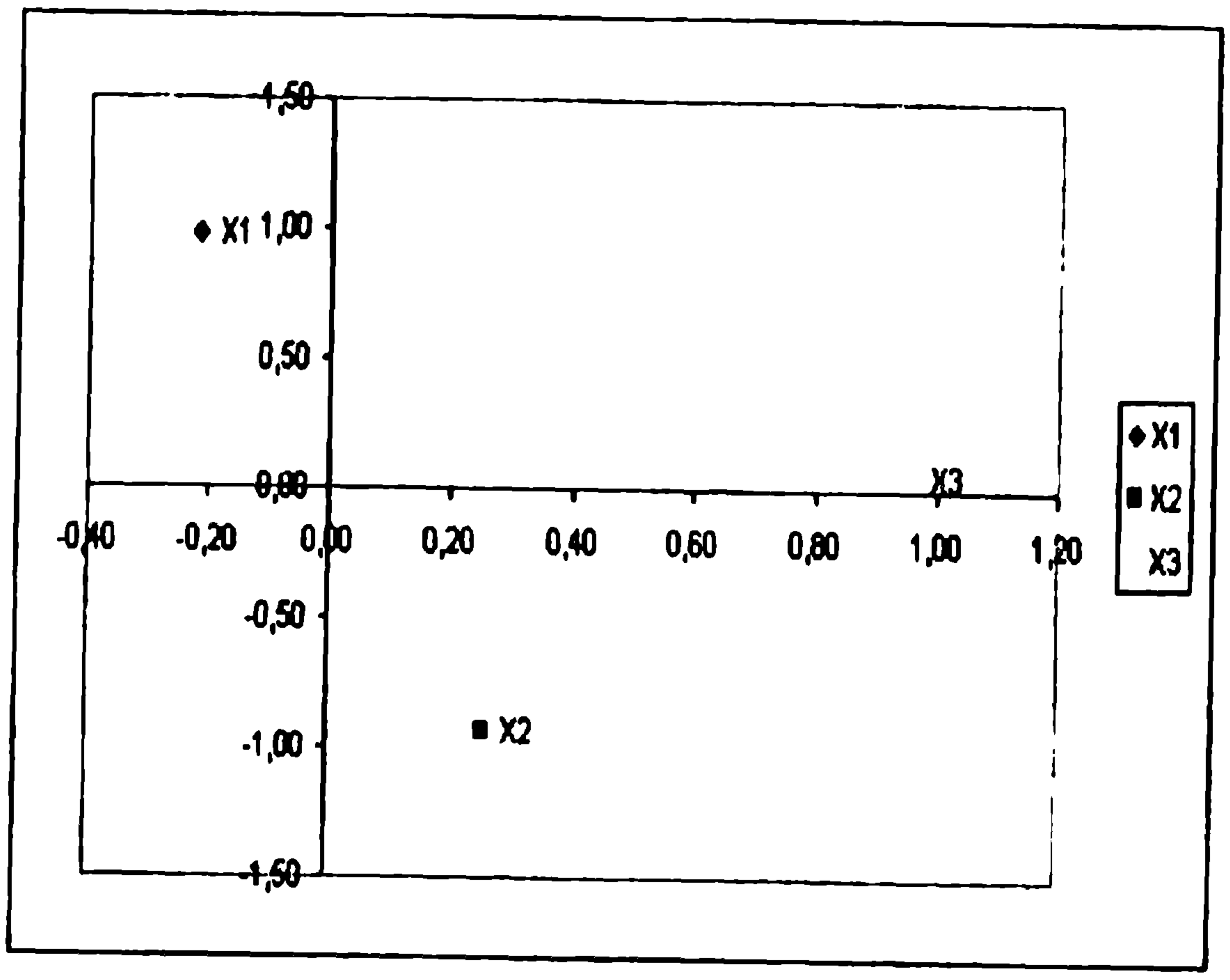
D√λ			
	14,62	0	
D√λ	0	6,952	0
	0	0	0,33

At			
	0,1	-0,04	-0,99
At	0,94	0,31	-0,11
	-0,32	-0,95	0,01

احدثيات المتغيرات في المحـــاور :

	X1	X2	X3
المحور1	-0,22	0,26	0,96
المحور2	0,97	-0,94	0,05
المحور3	0,02	0,14	0,00

إسقاط المتغيرات:



الفصل الثاني

.1.2

1.اختبار χ2

$$X^2 = \sum_{i=1}^{n} \sum_{j=1}^{p} \frac{(n_{ij} - (n_{i.}n_{.j}/n_{..}))^2}{n_{i.}n_{.j}/n_{..}} = 68.03$$

القيمة المجدولة χ2(0.95) مع درجة حرية 9 = 16,92.

القيمة المحسوبة أكبر من القيمة المجدولة إذا يوجد ارتباط بين السؤال الأول والثاني.

.2

$$d_{ij} = \frac{n_{ij} / n_{.j}}{n_{i.} / n_{..}}$$

i	Q1M1	Q1M2	Q1M3	Q1M4
Q2M1	**1,9910**	0,8440	0,7167	0,6588
Q2M2	0,5489	0,4266	**2,1758**	**1,6216**
Q2M3	**1,6319**	0,6204	0,8703	0,8571
Q2M4	0,3670	**1,9389**	0,6851	0,7518

نلاحظ تجاذب مابين كل من (Q2M1 ،Q1M1) ؛ (Q1M4، Q2M2) ؛ (Q2M2 ،Q1M3) ؛ (Q2M3 ،Q1M1)، (Q2M4 ،Q1M1) ؛ لأن قيمة مؤشر التجاذب و التنافر يفوق 1.

.3

H =

0,3884	0,2190	0,1563	0,1322
0,1290	0,1333	0,4326	0,3919
0,3336	0,2229	0,1989	0,1802
0,0913	0,5722	0,1906	0,1924

C =

0,2871	0,2289	0,2002	0,1795
0,2289	0,4428	0,2453	0,2314
0,2002	0,2453	0,2874	0,2627
0,1795	0,2314	0,2627	0,2405

.4

أ- القيمة الذاتية الباقيـة : $trace(C) = 1.25781$

$\lambda_0 = 1$ ، $\lambda_3 = (trace(C) - 1) - (0.1464 + 0.1113) = 0.0001$

ب- نسبة التمثيل

$$I = \frac{(0.1464 + 0.1113)}{0.25781} = 0.9996 \Rightarrow 99.96\%$$

5. المركبـات :

أ- مركبات الأسطـر :

$$y_{Ii}^{k} = \sum_{j} \frac{n_{ij}}{n_{i\bullet}} \sqrt{\frac{n_{\bullet\bullet}}{n_{\bullet j}}} a_{lk}^{j}$$

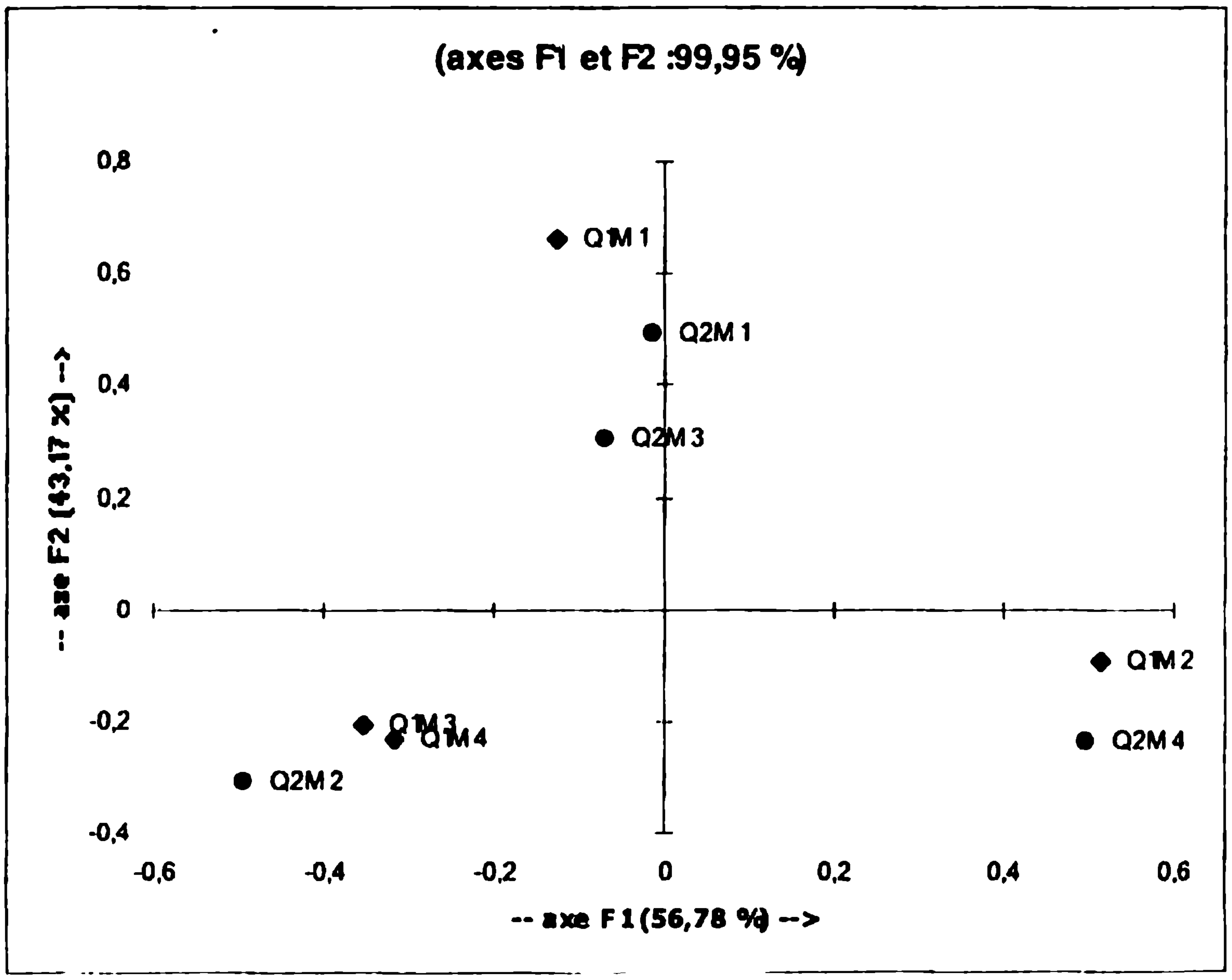
	المحور الأول	المحور الثاني
Q2M1	-0,0136	0,4957
Q2M2	-0,4929	-0,3084
Q2M3	-0,0708	0,3054
Q2M4	0,4956	-0,2357

ب- مركبات الأعمـدة :

$$y_{Jj}^{k} = \frac{1}{\sqrt{\lambda_k}} \sum_{i} \frac{n_{ij}}{n_{\bullet j}} y_{Ii}^{k}$$

	المحور الأول	المحور الثاني
Q1M1	-0,13	0,66
Q1M2	0,51	-0,09
Q1M3	-0,35	-0,20
Q1M4	-0,32	-0,23

-6



2.2. المساهمة المطلقة والنسبية للعناصر أو الأفراد الموجودة في الأعمدة :

- المساهمة المطلقة : $\alpha_j^K = \frac{n_{\bullet j}}{n_{\bullet\bullet}} \frac{(Y_{Jj}^k)^2}{\lambda_k} \times 100$

	$\alpha_j^1 = \frac{n_{\bullet j}}{n_{\bullet\bullet}} \frac{(Y_{Jj}^k)^2}{\lambda_k} \times 100$	$\alpha_j^2 = \frac{n_{\bullet j}}{n_{\bullet\bullet}} \frac{(Y_{Jj}^k)^2}{\lambda_k} \times 100$
Q1M1	2,15	78,07
Q1M2	62,36	2,70
Q1M3	21,15	9,20
Q1M4	14,35	10,03

- المساهمة النسبية : $\rho_j^K = \frac{(Y_{Jj}^k)^2}{\sum_{i=1}^{n} \frac{n_{\bullet\bullet}}{n_{i\bullet}} (\frac{n_{ij}}{n_{\bullet j}} - \frac{n_{i\bullet}}{n_{\bullet\bullet}})^2}$

	$\rho_j^1 = \frac{(Y_{Jj}^k)^2}{\sum_{i=1}^{n} \frac{n_{\bullet\bullet}}{n_{i\bullet}} (\frac{n_{ij}}{n_{\bullet j}} - \frac{n_{i\bullet}}{n_{\bullet\bullet}})^2}$	$\rho_j^2 = \frac{(Y_{Jj}^k)^2}{\sum_{i=1}^{n} \frac{n_{\bullet\bullet}}{n_{i\bullet}} (\frac{n_{ij}}{n_{\bullet j}} - \frac{n_{i\bullet}}{n_{\bullet\bullet}})^2}$
Q1M1	0,0350	0,9650
Q1M2	0,9681	0,0319
Q1M3	0,7512	0,2486
Q1M4	0,6527	0,3470

- المساهمة المطلقة والنسبية للعناصر أو الأفراد الموجودة في الأسطر:

- المساهمة المطلقة : $\alpha_i^K = \frac{n_{i\bullet}}{n_{\bullet\bullet}} \frac{(Y_{Ii}^k)^2}{\lambda_k}$

	$\alpha_i^1 = \frac{n_{i\bullet}}{n_{\bullet\bullet}} \frac{(Y_{li}^k)^2}{\lambda_k}$	$\alpha_i^2 = \frac{n_{i\bullet}}{n_{\bullet\bullet}} \frac{(Y_{li}^k)^2}{\lambda_k}$
Q2M1	0,02	42,62
Q2M2	46,51	23,94
Q2M3	0,73	17,77
Q2M4	52,74	15,68

- المساهمة النسبيـة : $$\rho_i^K = \frac{(Y_{li}^k)^2}{\sum_{j=1}^{p} \frac{n_{\bullet\bullet}}{n_{\bullet j}} (\frac{n_{ij}}{n_{i\bullet}} - \frac{n_{\bullet j}}{n_{\bullet\bullet}})^2}$$

	$\rho_i^1 = \frac{(Y_{li}^k)^2}{\sum_{j=1}^{p} \frac{n_{\bullet\bullet}}{n_{\bullet j}} (\frac{n_{ij}}{n_{i\bullet}} - \frac{n_{\bullet j}}{n_{\bullet\bullet}})^2}$	$\rho_i^2 = \frac{(Y_{li}^k)^2}{\sum_{j=1}^{p} \frac{n_{\bullet\bullet}}{n_{\bullet j}} (\frac{n_{ij}}{n_{i\bullet}} - \frac{n_{\bullet j}}{n_{\bullet\bullet}})^2}$
Q2M1	0,0008	0,9991
Q2M2	0,7187	0,2813
Q2M3	0,0510	0,9486
Q2M4	0,8156	0,1844

.3.2

.1

H=

0	**0**	**0,2886**	**0,2886**
0,3162	**0,25**	**0**	**0**
0,2581	**0,2041**	**0,2357**	**0**
0	**0,25**	**0**	**0,2886**
0,2581	**0**	**0,23570226**	**0,2357**
0	**0,25**	**0**	**0,2886**

0,2581	**0,2041**	**0,2357**	**0**
0	**0,25**	**0**	**0,2886**
0	**0,25**	**0,2886**	**0**
0,2236	**0,1767**	**0,2041**	**0,2041**

C=

0,3500	0,2240	0,2282	0,1065
0,2240	0,4271	0,2045	0,2526
0,2282	0,2045	0,3750	0,1806
0,1065	0,2526	0,1806	0,4306

.2 $\lambda_0 = 1$ ، $trace\ (C) = 1.5821$

- **نسبة التمثيل :** $I = \frac{(0.303 + 0.181)}{0.5821} = 0.83.14 \Rightarrow 83.14\%$

$$y_{li}^{k} = \sum_{j} \frac{n_{ij}}{n_{i\bullet}} \sqrt{\frac{n_{\bullet\bullet}}{n_{\bullet j}}} a_{lk}^{j}$$

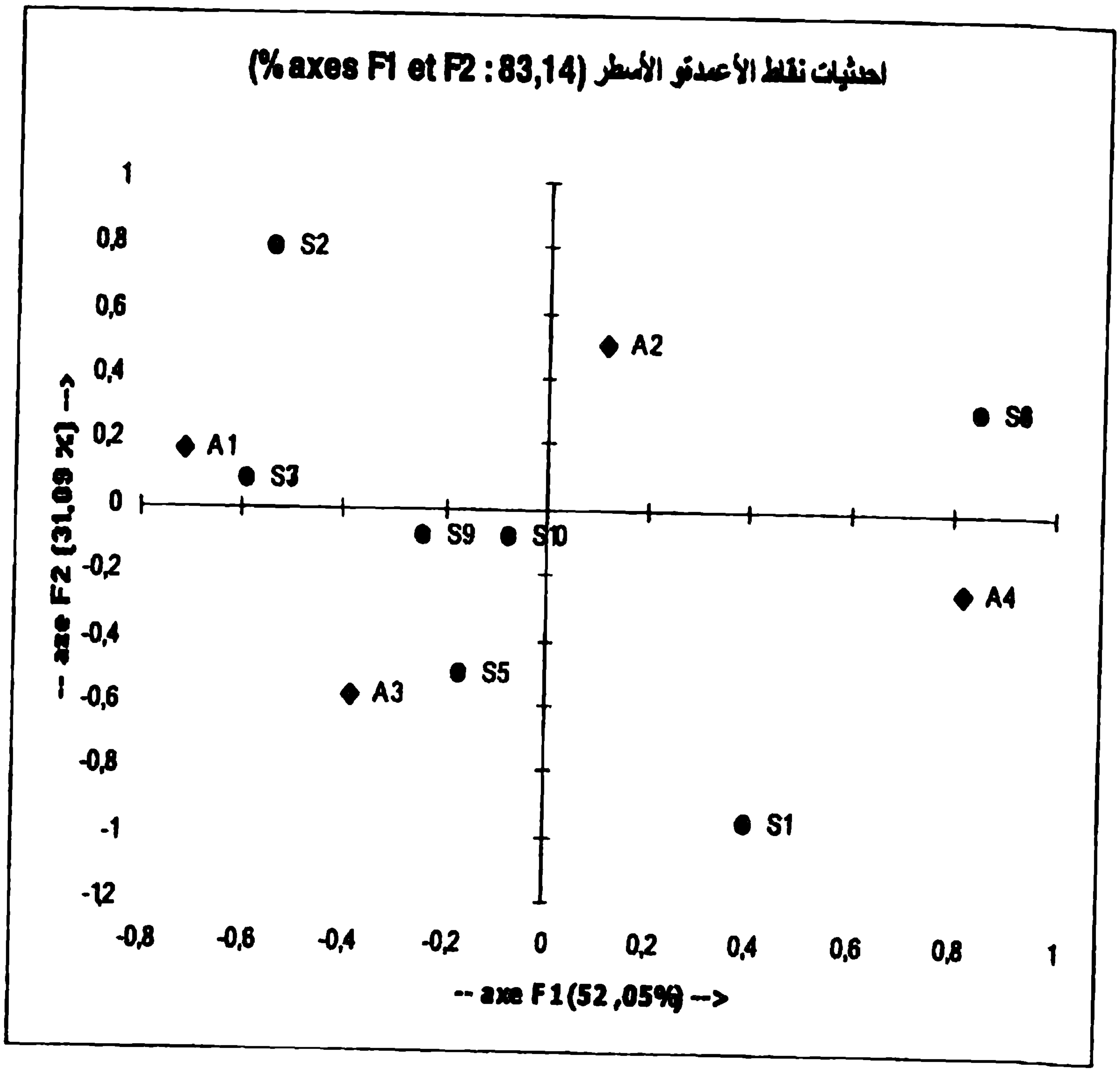
	المحــور الأول	المحور الثانــي
S1	0,3977	-0,9596
S2	-0,5410	0,7956
S3	-0,5927	0,0790
S4	0,8521	0,3107
S5	-0,1665	-0,5048
S6	0,8521	0,3107
S7	-0,5927	0,0790
S8	0,8521	0,3107
S9	-0,2417	-0,0839
S10	-0,0717	-0,0820

بالنسبة للأعمدة :

$$y_{Jj}^{k} = \frac{1}{\sqrt{\lambda_k}} \sum_i \frac{n_{ij}}{n_{\bullet j}} y_{Ii}^{k}$$

	المحور الأول	المحور الثاني
A1	-0,7133	0,1723
A2	0,1172	0,5050
A3	-0,3835	-0,5765
A4	0,8217	-0,2405

تمثيل الأسئلة والمواضيع :

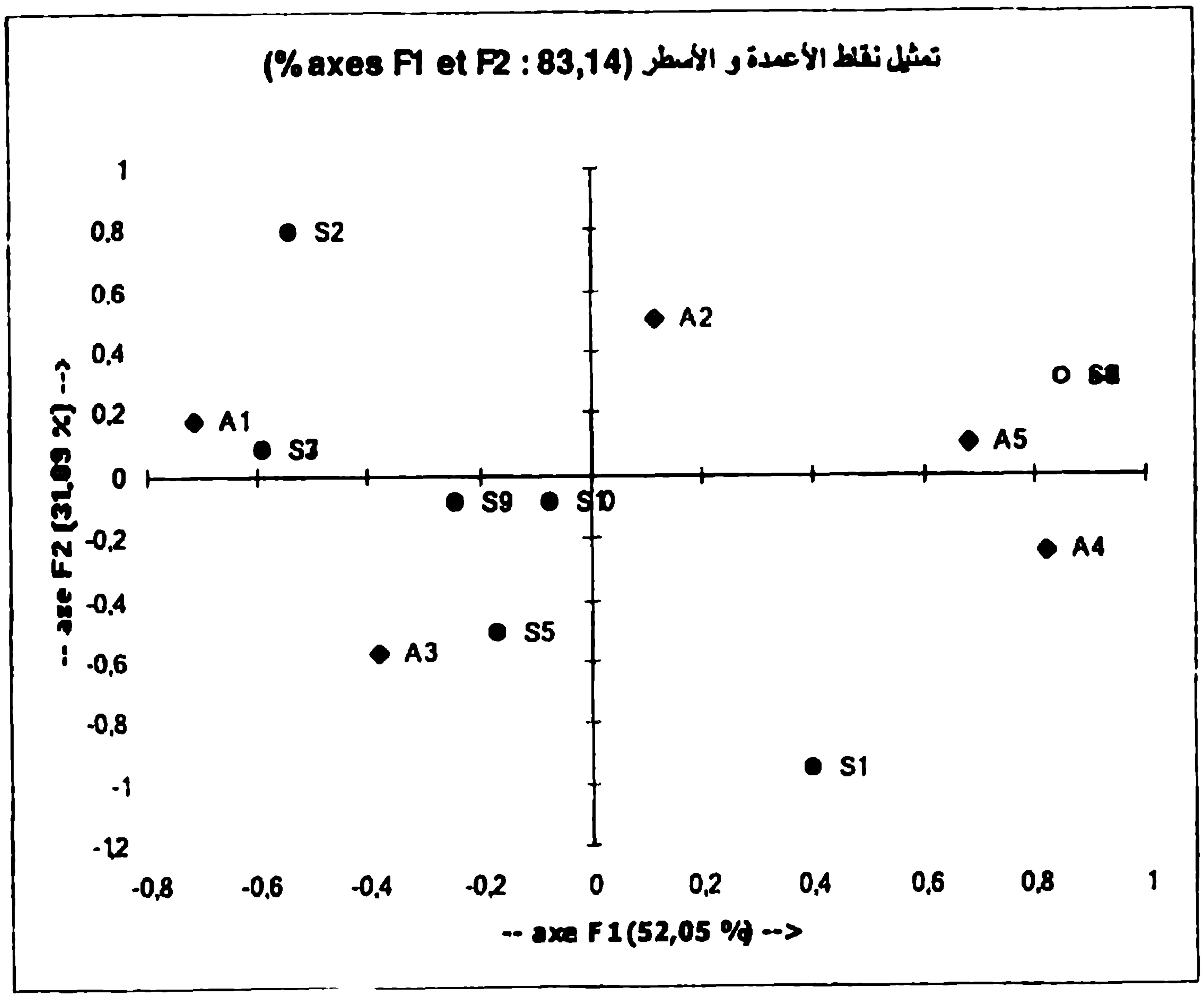


3- مركبات السؤال الخامس :

A5	
1	
1	
0	
1	
1	
1	
0	
1	
0	
0	
6	المجموع

$$Y_{Jc}^{K} = \frac{}{\sqrt{\lambda_k}} \sum_{i=1}^{n} \frac{n_{ic}}{n_{\bullet c}} Y_{Ii}^{k}$$

	المحور الأول	المحور الثاني
A5	0,6797	0,1031



مركبات الموضوع 11

					المجموع
S11	0	1	0	01	3

$$Y_{ls}^{K} = \frac{}{\sqrt{\lambda_k}} \sum_{j=1}^{p} \frac{n_{sj}}{n_{s\bullet}} Y_{lj}^{k}$$

	المحور الأول	المحور الثاني
s11	0,8521	0,3107

الفصل الثالث

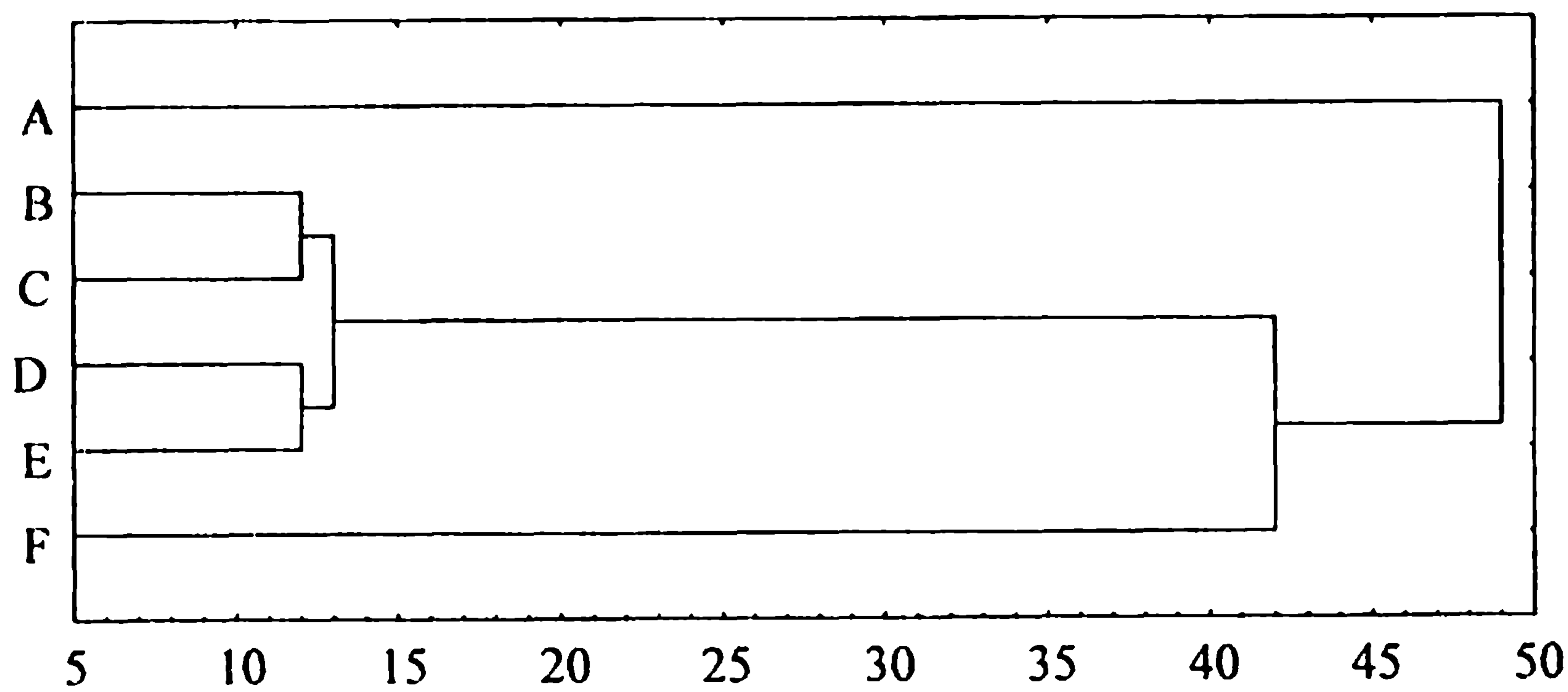
.1.3

1. مصفوفة المسافات للأفراد مربعة :

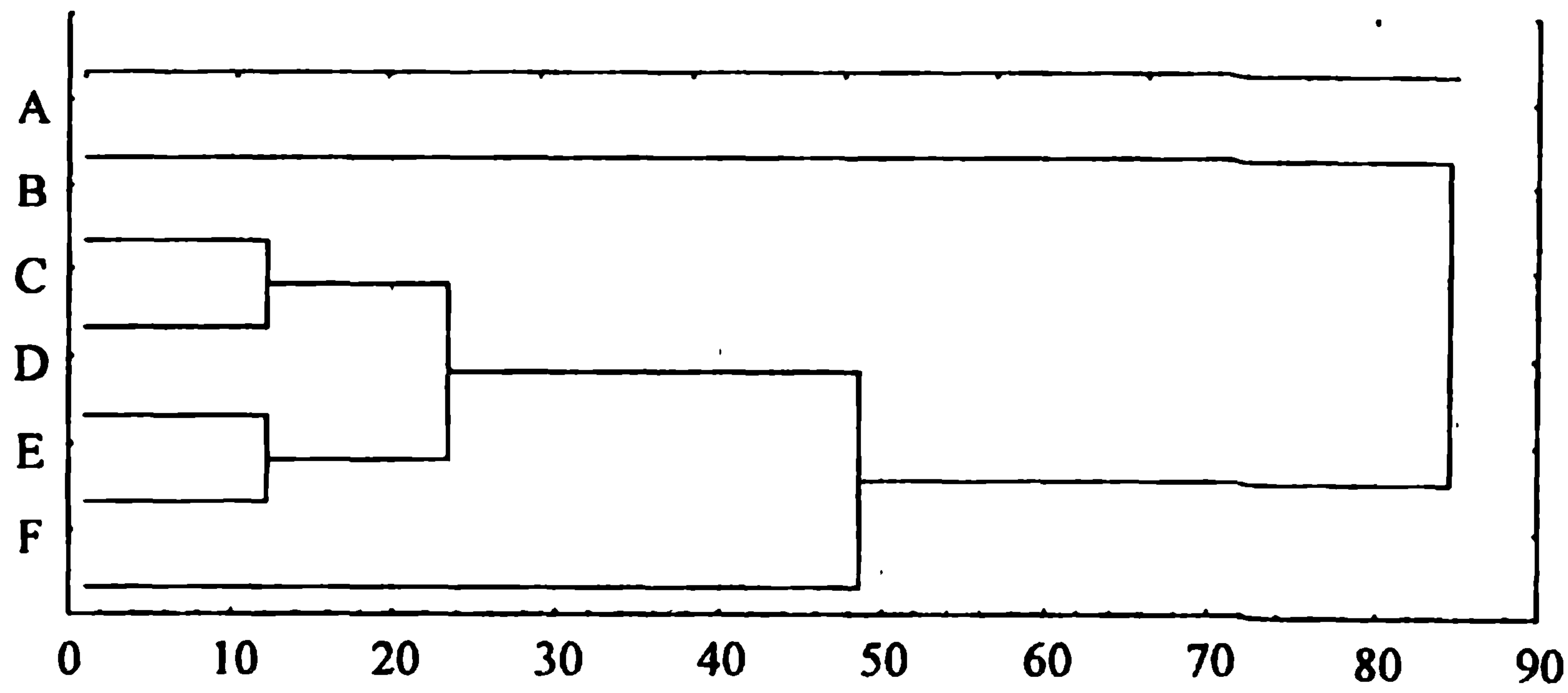
Observation	A	B	C	D	E	F
A		101	117	59	50	116
B	101		16	22	39	49
C	117	16		16	49	63
D	59	22	16		13	55
E	50	39	49	13		46
F	116	49	63	55	46	

.2

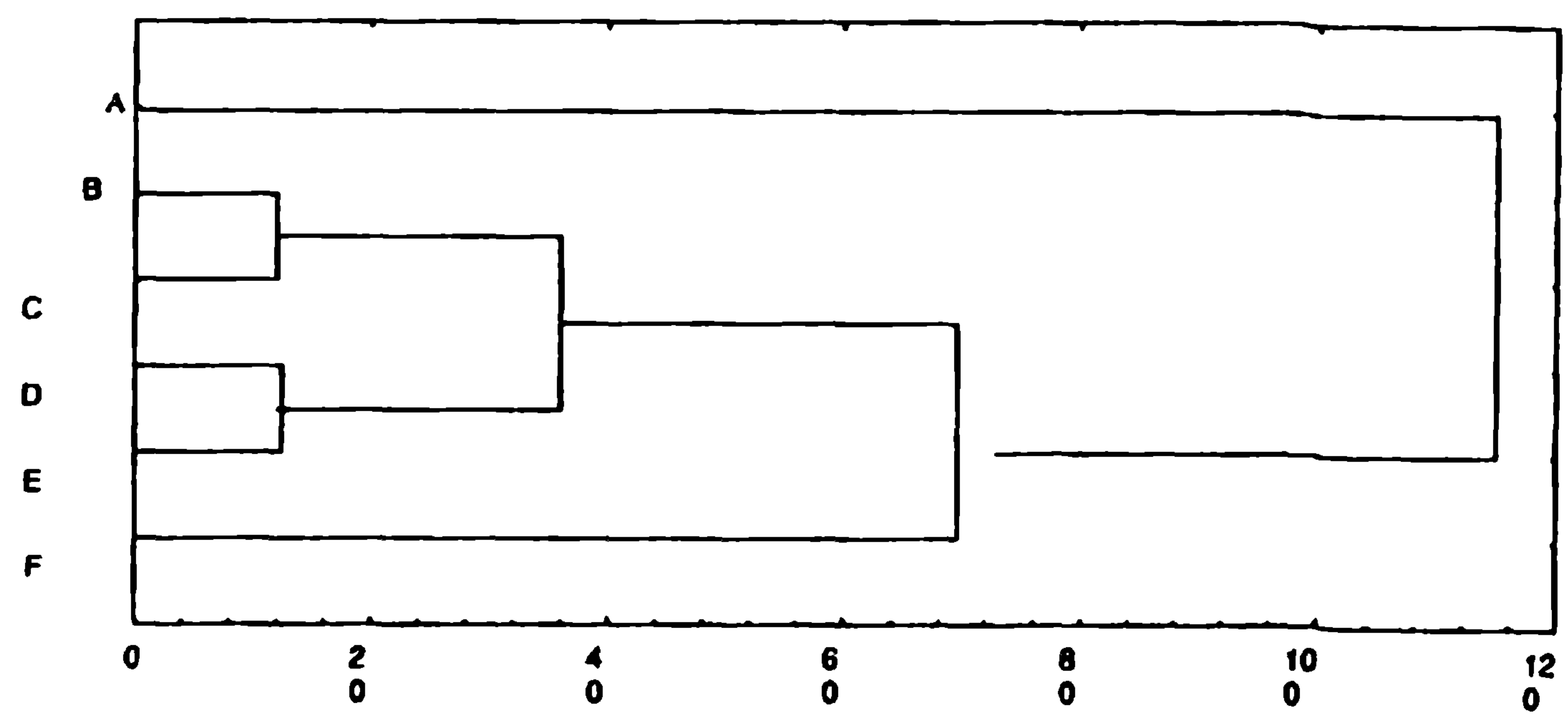
أ- الدندوقرام للمؤشر التجميع لأدنى مسافة :



ب- الدندوقرام للمؤشر التجميع لمتوسط المسافة



ج- الدندوقرام للمؤشر التجميع لـ WARD



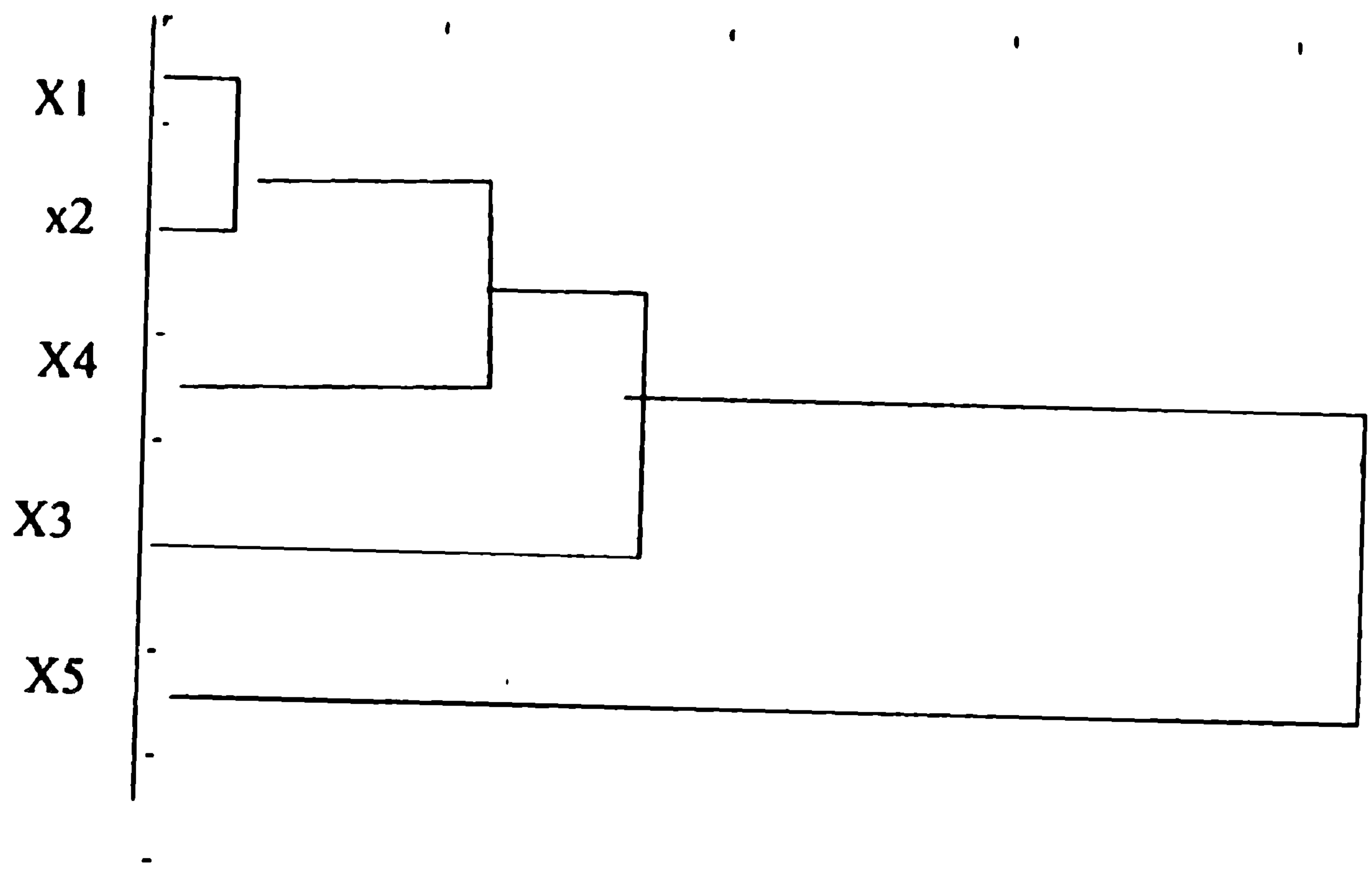
.2.3

1. جدول المسافات للمتغيـــرات

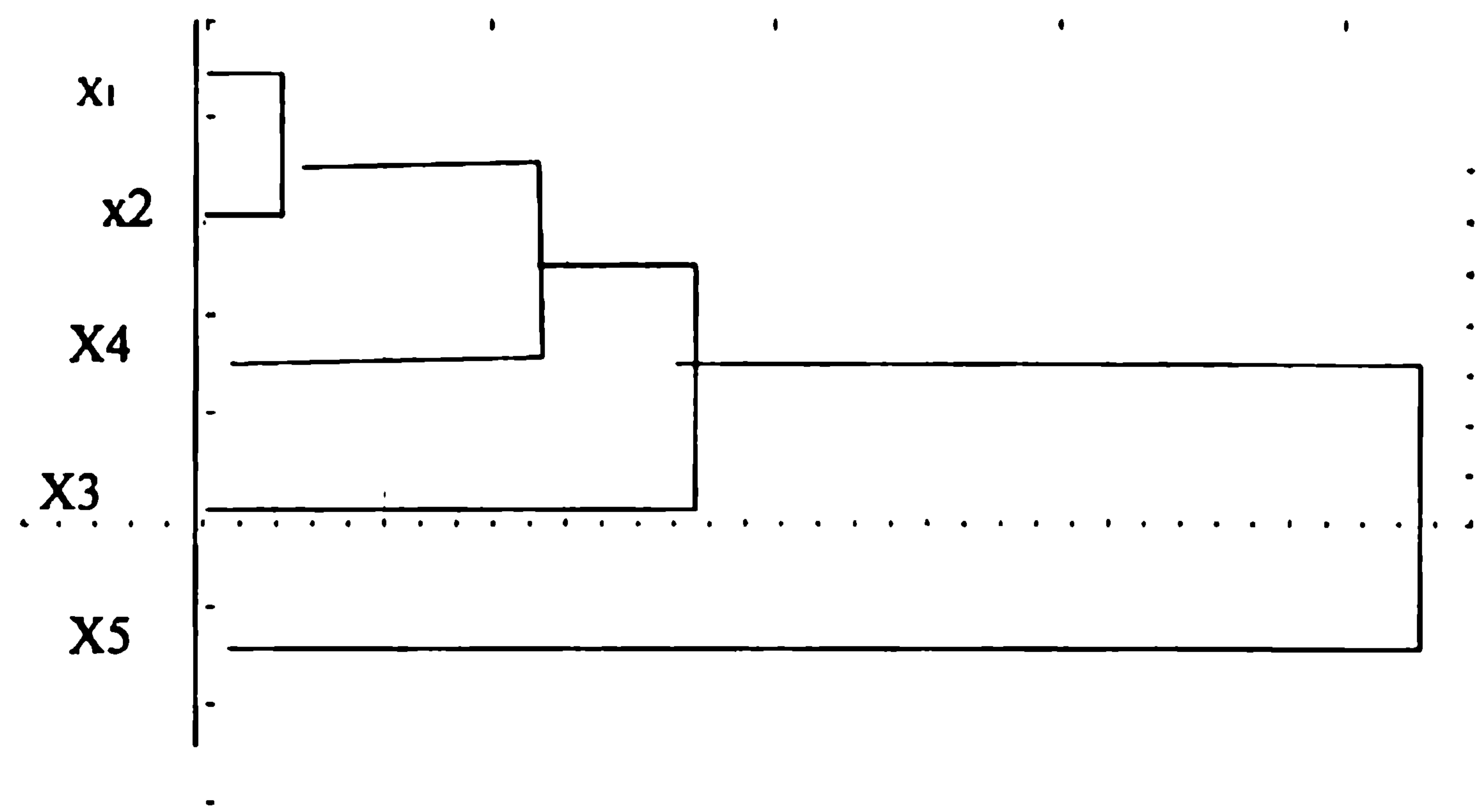
	X1	X2	X3	X4	X5
X1		22	45	59	110
X2	22		85	39	120
X3	45	85		100	131
X4	59	39	100		233
X5	110	120	131	233	

.2

أ- تصنيف المتغيرات باستعمال أدنى مسافة :



ب- تصنيف المتغيرات باستعمال طريقة WARD

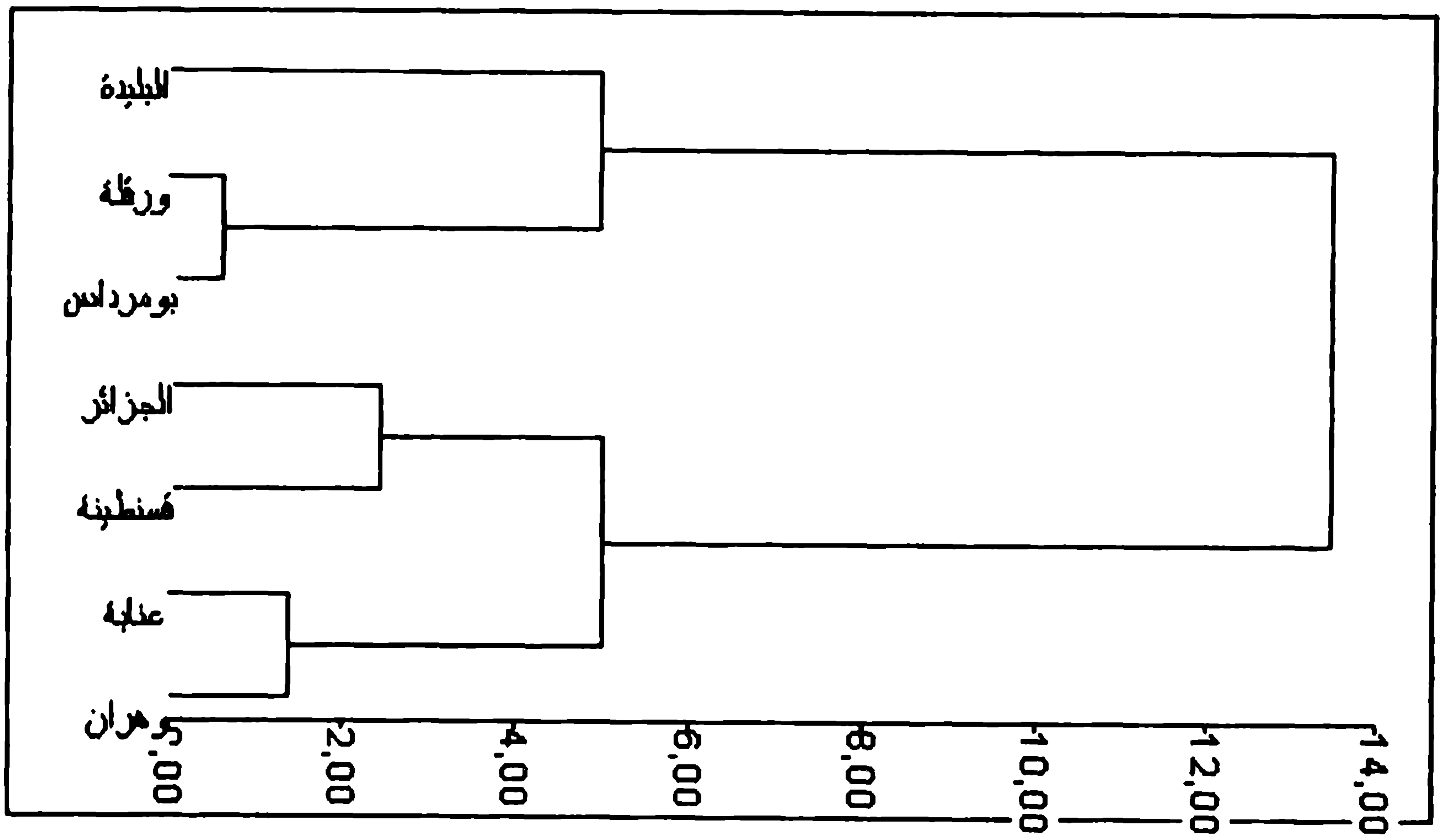


.3.3

1. مصفوفة المسافات :

	الجزائر	ورقلة	البليدة	بومرداس	قسنطينة	عنابة	وهران
الجزائر	0	3,55	4,38	4,57	2,22	2,45	3,16
ورقلة	3,55	0	2,60	1,12	2,64	2,59	1,14
البليدة	4,38	2,60	0	2,97	4,47	4,27	2,99
بومرداس	4,57	1,12	2,97	0	3,56	3,50	1,96
قسنطينة	2,22	2,64	4,47	3,56	0	2,20	2,64
عنابة	2,45	2,59	4,27	3,50	2,20	0	1,67
وهران	3,16	1,14	2,99	1,96	2,64	1,67	0

2. التصنيف عن طريق طريقة WARD



الـملحق الأول

التحويـلات في الـمصفـوفات

1. القيم الذاتية والأشعة الذاتيـــة

يعطـى في الكـثير مـن المـسائل في العلـوم والرياضـيات تطبيـق خطي $f : V \to V : f$ ويكون من الأساسي تعيين تلك الأعداد القياسية λ ؛ بحيث تكون للمعادلة $fx = \lambda x$ حل غير صفري.

1.1. **القيم الذاتية** : إذا كانت A مصفوفة من النمط $n \times n$ ؛ يسمى شعاعاً ذاتيا للمصفوفة A، الشعاع الغير صفري X في R^n ؛ إذا كانت AX مضاعفا قياسيا للشعاع X أي أن : $AX = \lambda X$

حيث : تسمى λ القيمة الذاتية للمصفوفة A، وX الشعاع ذاتي للعدد λ.

مثال : الشعاع $X = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ هو شعاع ذاتي للمصفوفة $A = \begin{pmatrix} 3 & 0 \\ 8 & -1 \end{pmatrix}$ ؛ تقابله القيمة الذاتية $\lambda = 3$ وبالتالـي : $AX = \begin{pmatrix} 3 & 0 \\ 8 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 3 \\ 6 \end{pmatrix} = 3X$

لإيجاد القيمة الذاتية لمصفوفة A من النمط $n \times n$ نكتب :

$AX = \lambda X$ على الشكل التالي : $AX = \lambda I_n X$ ، دائماً $A \neq 0$

أو : $(\lambda I_n - A)X = 0$

ولكي تكون λ قيمة ذاتية، يجب أن يكون لهذه المعادلة حل غير صفري، ولكي يكون لهذه المعادلة حل غير صفري ؛ يجب أن تتوفر فيها مايلي : $\det(\lambda I_n - A)X = 0$

تسمى هذه المعادلة بالمعادلة المميزة للمصفوفة A، والكثير الحدود $P_A(x) = \det(A - XI_n)$ بالكثير الحدود المميز لـ A.

مثال : أوجد القيم الذاتية للمصفوفة $A = \begin{pmatrix} 3 & 2 \\ -1 & 0 \end{pmatrix}$

الحل :

$$\lambda I - A = \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 3 & 2 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} \lambda - 3 & -2 \\ 1 & \lambda \end{pmatrix}$$

وأن : $\det(\lambda I - A) = \det \begin{pmatrix} \lambda - 3 & 2 \\ 1 & \lambda \end{pmatrix} = \lambda^2 - 3\lambda + 2 = 0$

حلول هذه المعادلة هي : $\lambda=1$، $\lambda=2$، وهي القيم الذاتية للمصفوفة A.

مثال : أوجد القيم الذاتية للمصفوفة $A = \begin{pmatrix} -2 & -1 \\ 5 & 2 \end{pmatrix}$

الحل:

$$\lambda I - A = \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} -2 & -1 \\ 5 & 2 \end{pmatrix} = \begin{pmatrix} \lambda + 2 & 1 \\ -5 & \lambda - 2 \end{pmatrix}$$

وأن : $\det(\lambda I - A) = \det \begin{pmatrix} \lambda + 2 & 1 \\ -5 & \lambda - 2 \end{pmatrix} = \lambda^2 + 1 = 0$

توجد حلول هذه المعادلة في مجموعة الأعداد المركبة C ؛ مع العلم أن $\lambda^2 = -1$ ؛ إذا الحلول هي $\lambda = -i$ ، $\lambda = i$.

2.1. الأشعة الذاتية : إن الأشعة الذاتية للمصفوفة A المناظرة للقيمة ذاتية λ هي الأشعة الذاتية التي تحقق $AX=\lambda X$ بطريقة مكافئة؛ الأشعة الذاتية المناظرة للقيمة الذاتية λ هي الأشعة الصفرية في الفضاء الحلول للمعادلة $(\lambda I_n - A)X = 0$ ؛ ويسمى بفضاء الحلول للفضاء الذاتي للمصفوفة A المناظرة للقيمة الذاتية λ.

مثال : أوجد الأساسات الذاتية للمصفوفة التالية :

$$A = \begin{pmatrix} 3 & -2 & 0 \\ -2 & 3 & 0 \\ 0 & 0 & 5 \end{pmatrix}$$

الحل :

1. نقوم بحساب القيم الذاتية :

$$\lambda I - A = \lambda \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \begin{pmatrix} 3 & -2 & 0 \\ -2 & 3 & 0 \\ 0 & 0 & 5 \end{pmatrix} = \begin{pmatrix} \lambda-3 & +2 & 0 \\ -2 & \lambda-3 & 0 \\ 0 & 0 & \lambda-5 \end{pmatrix}$$

إذا :

$$\det(\lambda I - A) = (\lambda-5)^2(\lambda-1) = 0$$
$$\lambda = 5, \lambda = 1$$

2. إيجاد الأشعة الذاتية :

بالتعريف $X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$

نحصل على الأشعةً الذاتية للمصفوفة A المقابلة للقيمة λ إذا وفقط إذا كان X حلا غير تافه للمعادلة $(\lambda I - A)X = 0$

أي المعادلة :

$$\begin{pmatrix} \lambda-3 & +2 & 0 \\ -2 & \lambda-3 & 0 \\ 0 & 0 & \lambda-5 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

إذا كان λ=5 ؛ فإن :

$$\begin{pmatrix} 2 & 2 & 0 \\ -2 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

حل هذا النظام هو : $x_1 = -S, x_2 = S, x_3 = t$

إذا فإن الأشعة الذاتية للمصفوفة A المقابلة للقيمة λ=5 هي الأشعة غير الصفرية التالية :

$$X = \begin{pmatrix} -S \\ S \\ t \end{pmatrix} = \begin{pmatrix} -S \\ S \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ t \end{pmatrix} = S\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + t\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

حيث أن : $\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$ و $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ غير مرتبطة خطيا فإنهما يكوننا أساسا للفضاء الذاتي المناظر للقيمة الذاتية λ = 5

- إذا كان λ=1 ؛ فإن :

$$\begin{pmatrix} -2 & 2 & 0 \\ -2 & -2 & 0 \\ 0 & 0 & -4 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

وعليه حل هذا النظام هو: $x_1 = t, x_2 = t, x_3 = 0$

إذا فإن الأشعة الذاتية للمصفوفة A المناظرة لقيمة $\lambda=1$ هي الأشعة غير الصفرية التالية :

$$X=\begin{pmatrix} -t \\ t \\ 0 \end{pmatrix}=t\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

إذن الشعاع هو $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ أساس للفضاء الذاتي للقيمة الذاتية $\lambda=1$

2. التحويلات إلى شكل قطري والأشعة الذاتية

نظرية 1: ليكن f تطبيق من $f: E\rightarrow E$ ؛ حيث E هي فضاء شعاعي منتهي، التطبيق f يمكن تمثله بمصفوفة قطرية إذا وفقط إذا كان لـ E أساس مكون من الأشعة الذاتية لـ f، وفي هذه الحالة العناصر القطرية لهذه المصفوفة هي القيمة الذاتية المناظرة.

أي بصفة بديلة : تكون المصفوفة A ذات النمط $n\times n$ مصفوفة قطرية إذا وفقط إذا كانت لـ A أشعة ذاتية مستقلة خطيا عددها n، وفي هذه الحالة العناصر القطرية للمصفوفة القطرية D هي القيم الذاتية المناظرة للأشعة الذاتية A .

إذا افترضنا في النظرية السابقة أن P هي المصفوفة التي أعمدتها الأشعة الذاتية المستقلة (التي عددها n)، فإن $D=P^{-1}AP$.

مثال : أوجد المصفوفة P التي تحول إلى صورة قطرية المصفوفة :

$$A=\begin{pmatrix} 3 & -2 & 0 \\ -2 & 3 & 0 \\ 0 & 0 & 5 \end{pmatrix}$$

الحل : القيم الذاتية المرافقة للمصفوفة A هي λ=1 ، λ=5.

الأشعة الذاتية المناظرة للقيمة الذاتية λ=5 هي :

$$\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \text{ و } \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

والشعاع الذاتي المرافق للقيمة الذاتية λ=1 : $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$

إذا : $P=\begin{pmatrix} -1 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$

نحول A إلى صورة قطرية كالآتي :

$$P^{-1}AP=\begin{pmatrix} -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}\begin{pmatrix} 3 & -2 & 0 \\ -2 & 3 & 0 \\ 0 & 0 & 5 \end{pmatrix}\begin{pmatrix} -1 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}=\begin{pmatrix} 5 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

1.2. كثيرات الحدود للمصفوفات والتطبيقات الخطية : نعتبر كثير الحدود f(λ) على المجال K :

$$f(\lambda)=a_n\lambda^n+\ldots+a_1\lambda+a_0I$$

حيث: I هي المصفوفة الأحادية، وبوجه خاص نقول أن A جذر لكثير الحدود f(λ) ، إذا كانت $f(A)=0$

نظرية كايلي - هاملتون (Cayley Hamilton) : لتكن $A \in M_n(K)$ و g(λ) كثير حدود المميز للمصفوفة A، فإن $g(A)=0$.

2.2. أدنى كثير الحدود (Polynome Minimal)

نظرية 2: أدنى كثير حدود $P_A(x)$ لـ A، هو كل كثير حدود يقبل القسمة على A، وبوجه خاص فإن كثير الحدود المميز $\Lambda(x)$ للمصفوفة A تقبل القسمة على $P_A(x)$.

مثال : أوجد أدنى كثير حدود لـ$P_A(x)$ للمصفوفة.

$$A=\begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -2 & 4 \end{pmatrix}$$

الحل : $P_A(x)=\det(A-XI_n)$

$$\Delta(\lambda)=\begin{pmatrix} \lambda-2 & -1 & 0 & 0 \\ 0 & \lambda-2 & 0 & 0 \\ 0 & 0 & \lambda-1 & -1 \\ 0 & 0 & -2 & \lambda-4 \end{pmatrix}$$

$$\Delta(\lambda)=(\lambda-3)(\lambda-2)^3$$

أدنى كثير حدود لـ A هو أحد الكثيرات التالية :

$$f(x)=(x-3)(x-2),\ g(x)=(x-3)(x-2)^2,\ h(x)=(x-3)(x-2)^3$$

تحليل المعطيات

$$f(A)=(A-3I)(A-2I)=\begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -2 & 1 \\ 0 & 0 & 2 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & -2 & 1 \end{pmatrix} \neq 0$$

$$f(A)=(A-3I)(A-2I)^2=\begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -2 & 1 \\ 0 & 0 & 2 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & -2 & 1 \end{pmatrix}^2 = 0$$

إذن $g(x)=(x-3)(x-2)$ هو أدنى كثير لـ A.

الـملحق الثاني

تطبيق طريقة تحليل المعطيات بالمركبات

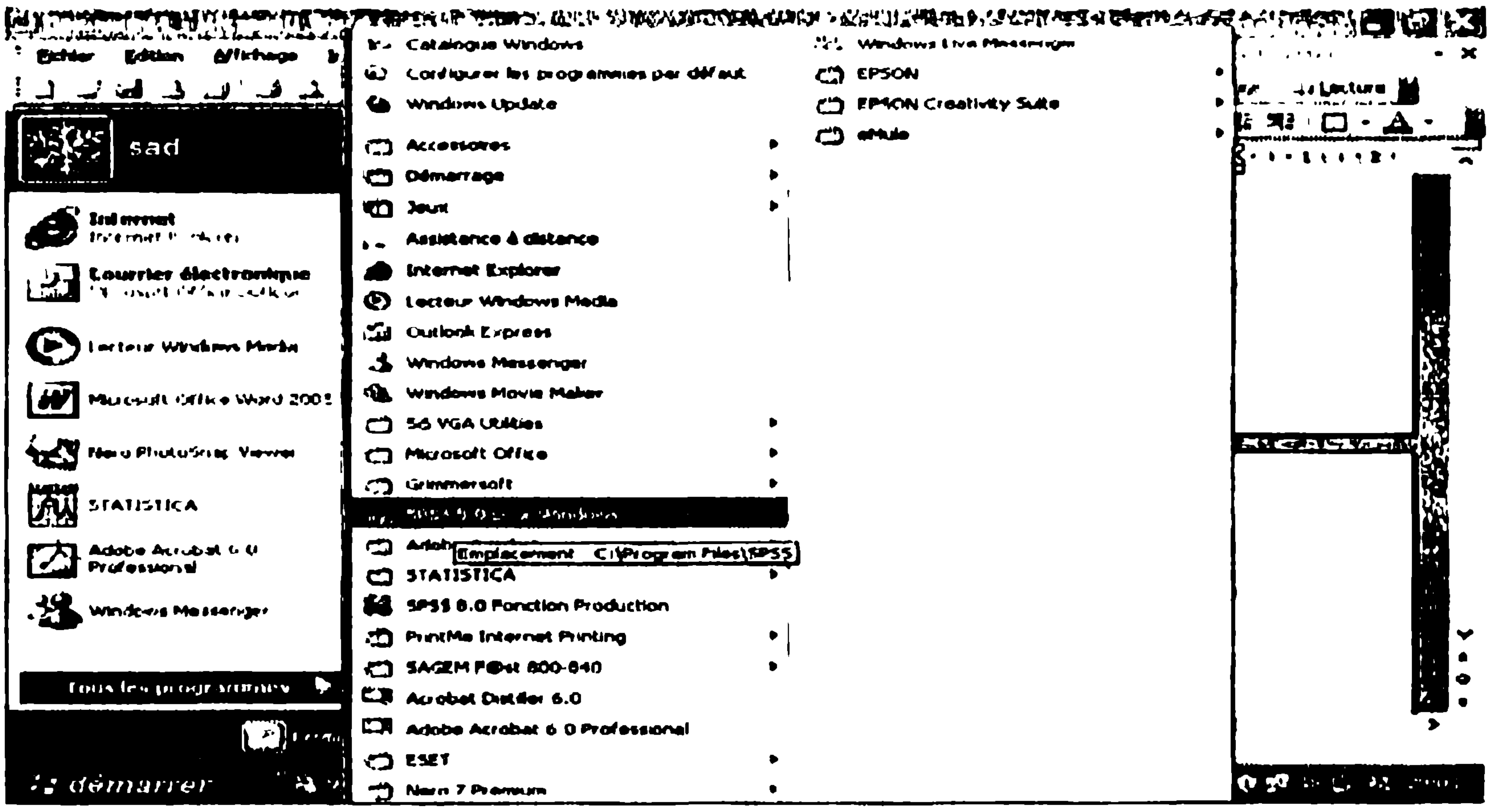
الرئيسية باستعمال SPSS

مثـــال :

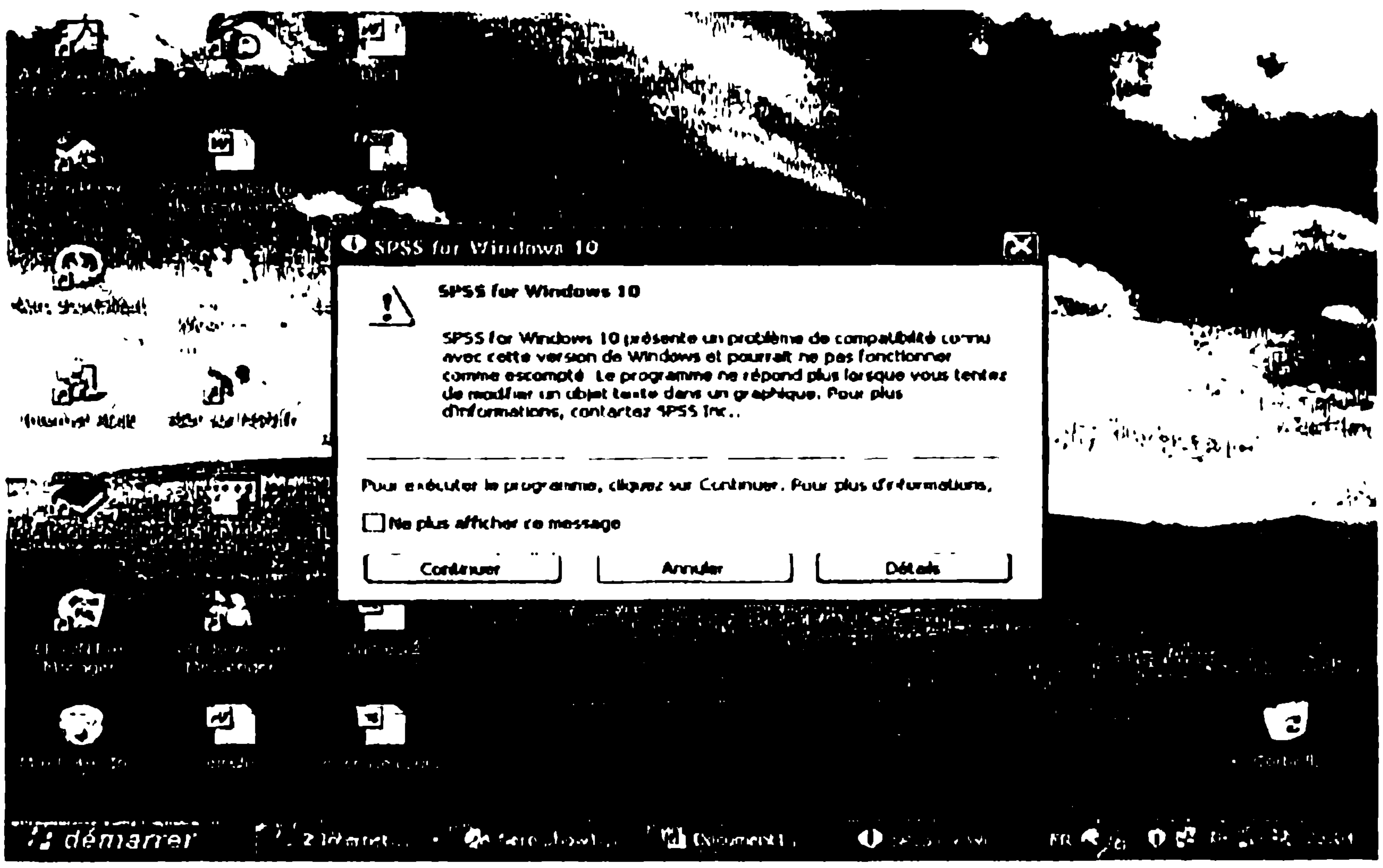
i	X1	X2	X3
a	1	2	4
b	2	3	6
c	5	4	8
d	3	5	10

- تحليل الجدول باستعمال طريقة ACP مع أخذ وحدة قياس المصفوفة الأحادية.

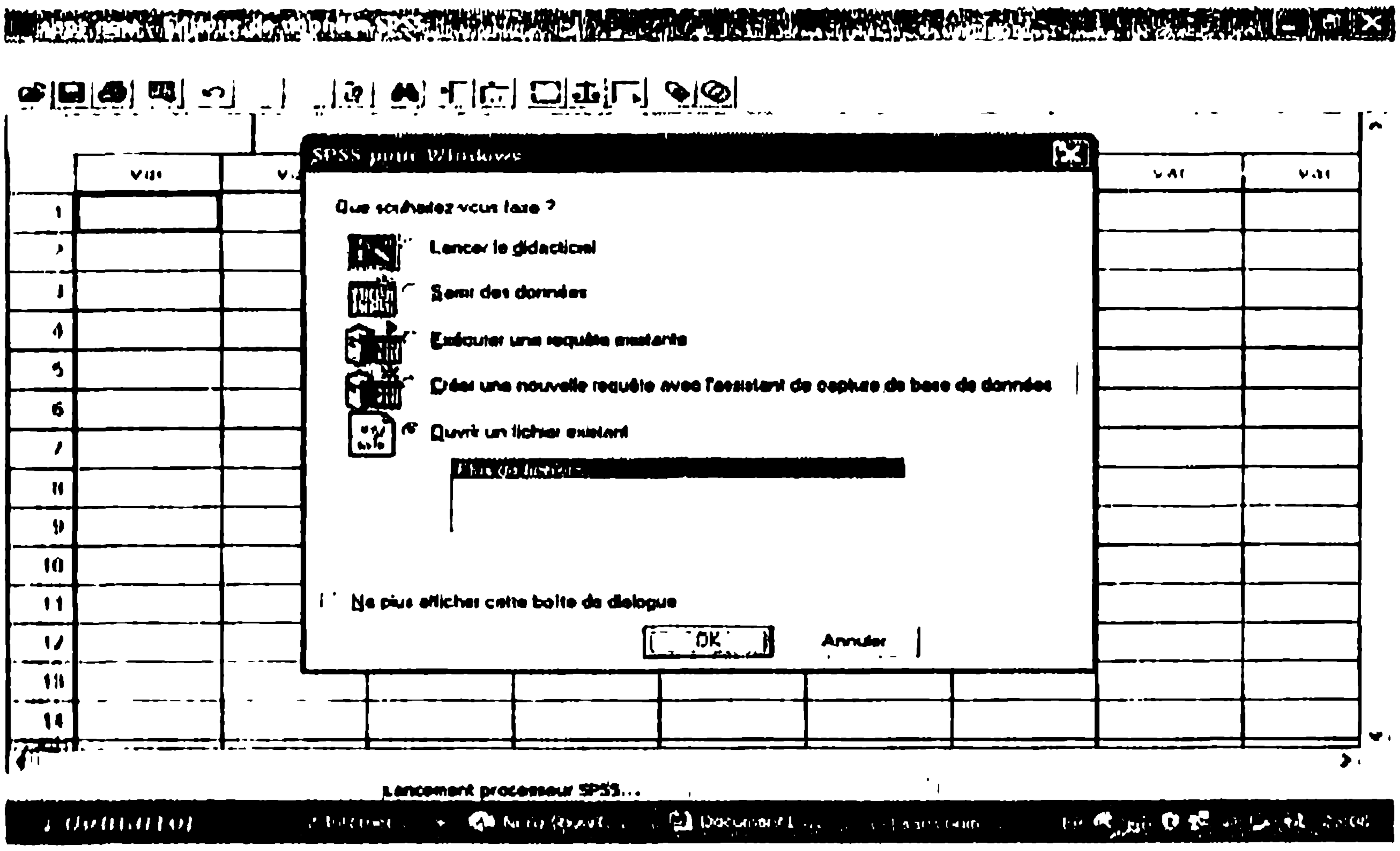
1. بعد تثبيت برنامجSPPS 8.0 ، نذهب إلى SPSS 8.0 pour Windows كما يلي :



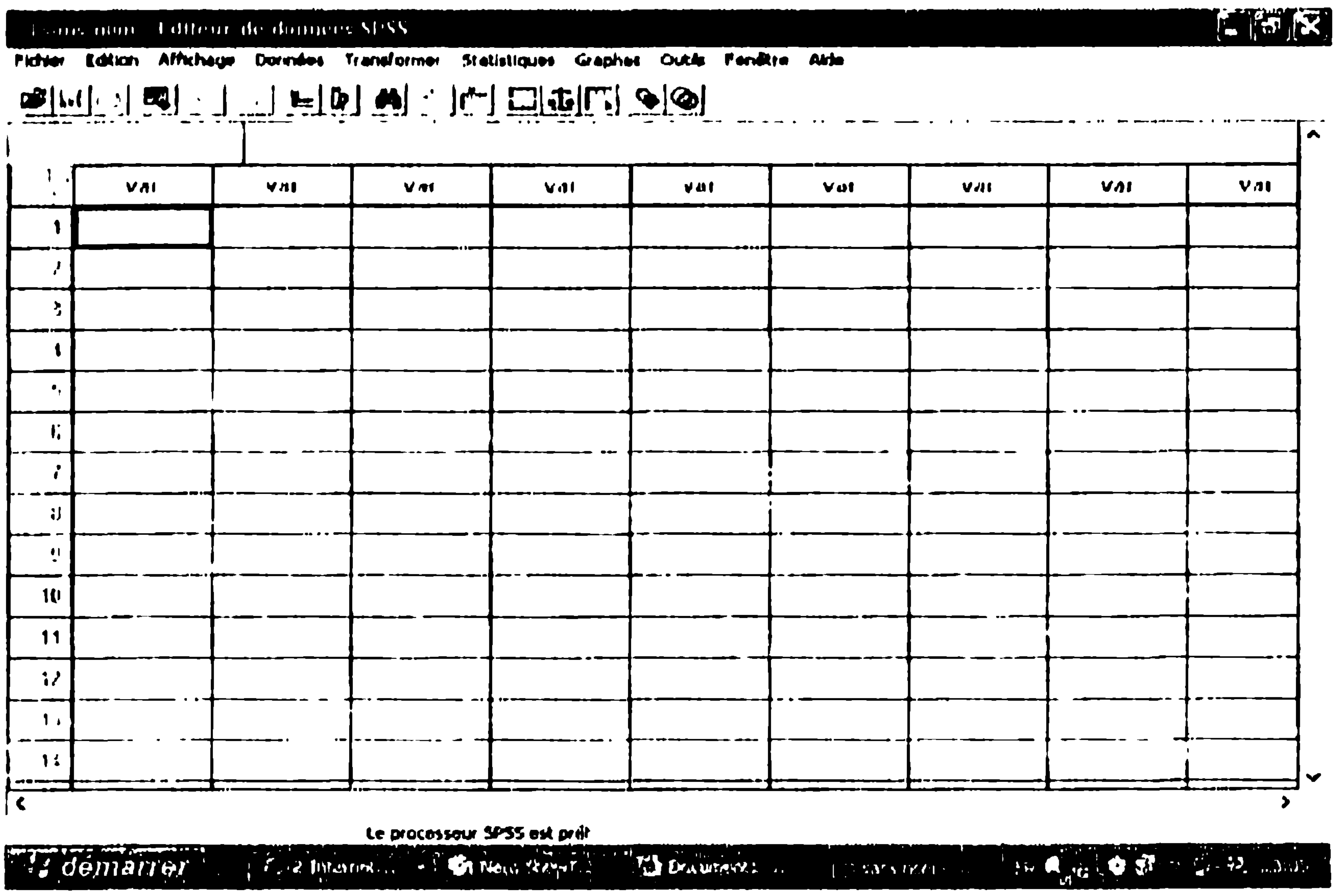
2. وعند الضغط نحصل على الشاشة التاليـة :



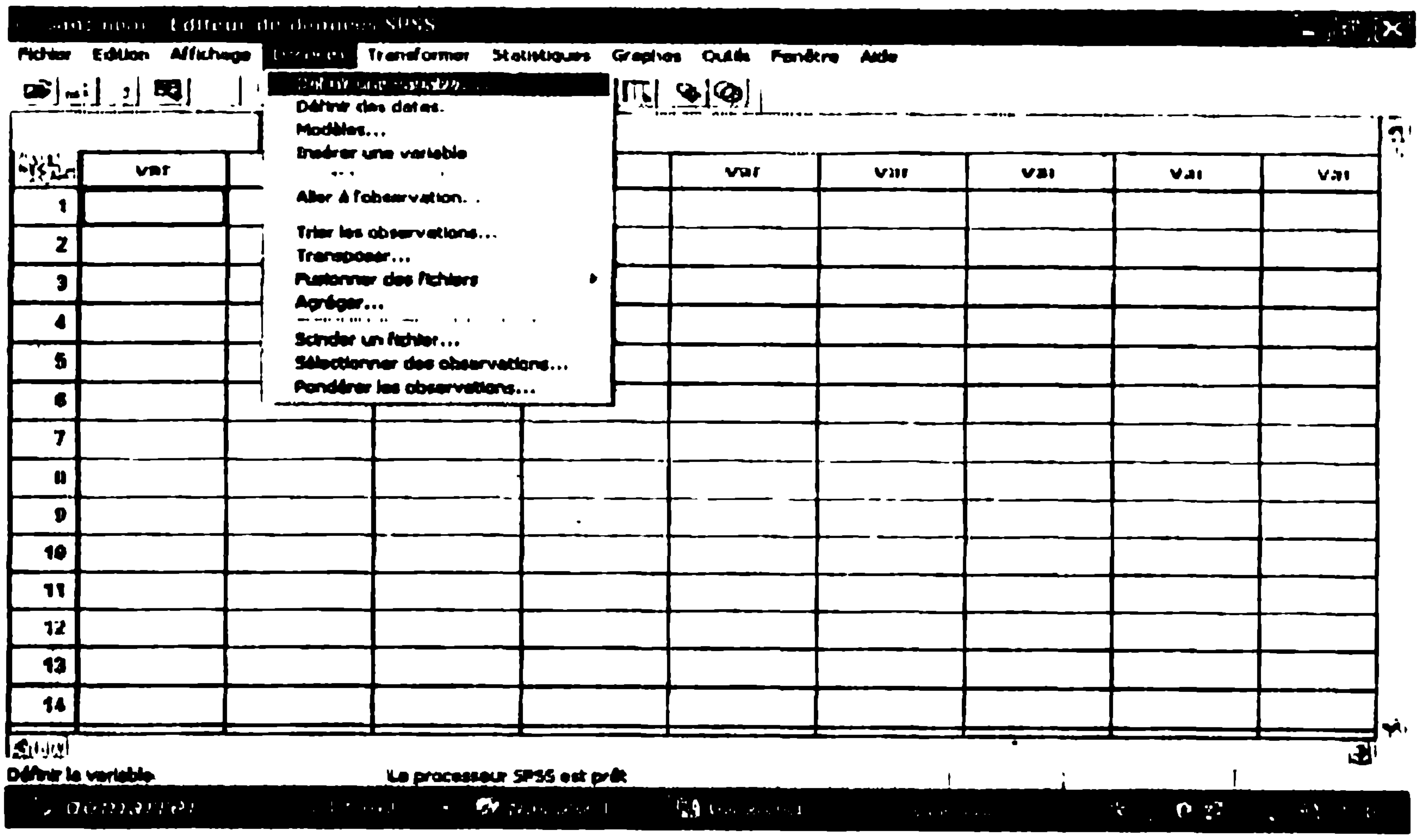
3. نحصل على الشاشة التالية بالضغط على Continuer :



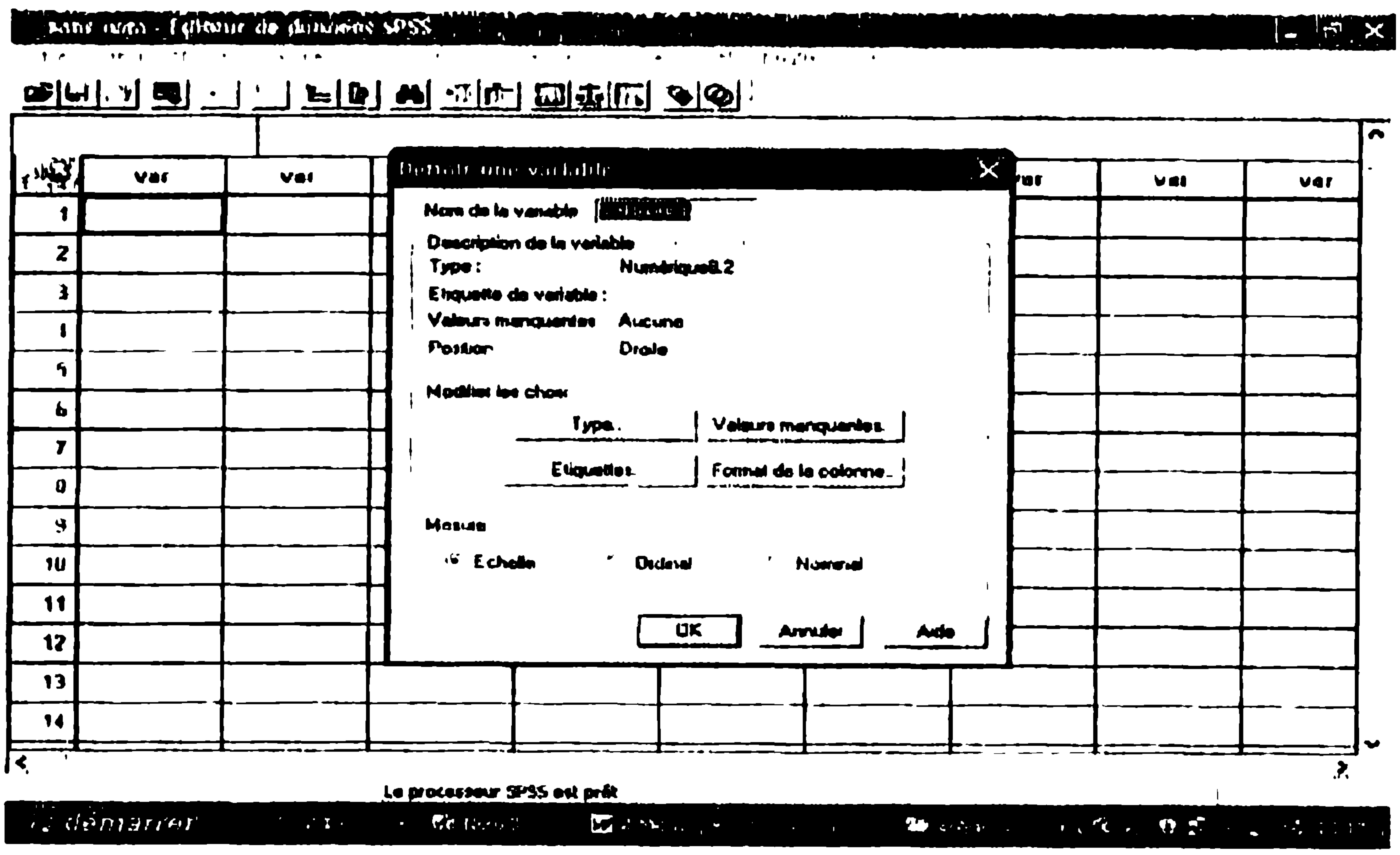
4. نقوم بالضغط على Annuler ، نحصل على ما يلي :



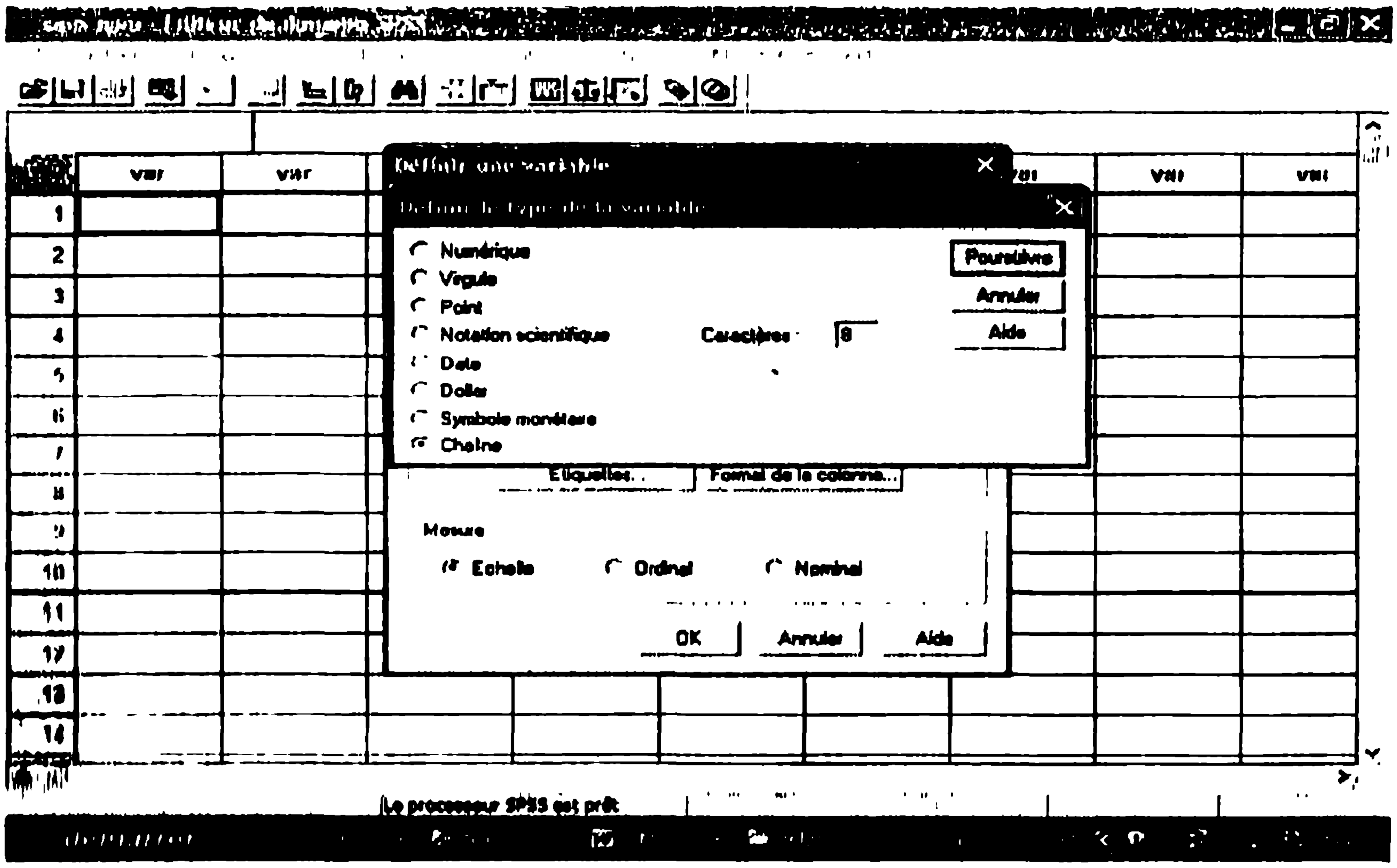
5. نذهب إلى Données، ونضغط على **Définir une Variable** :



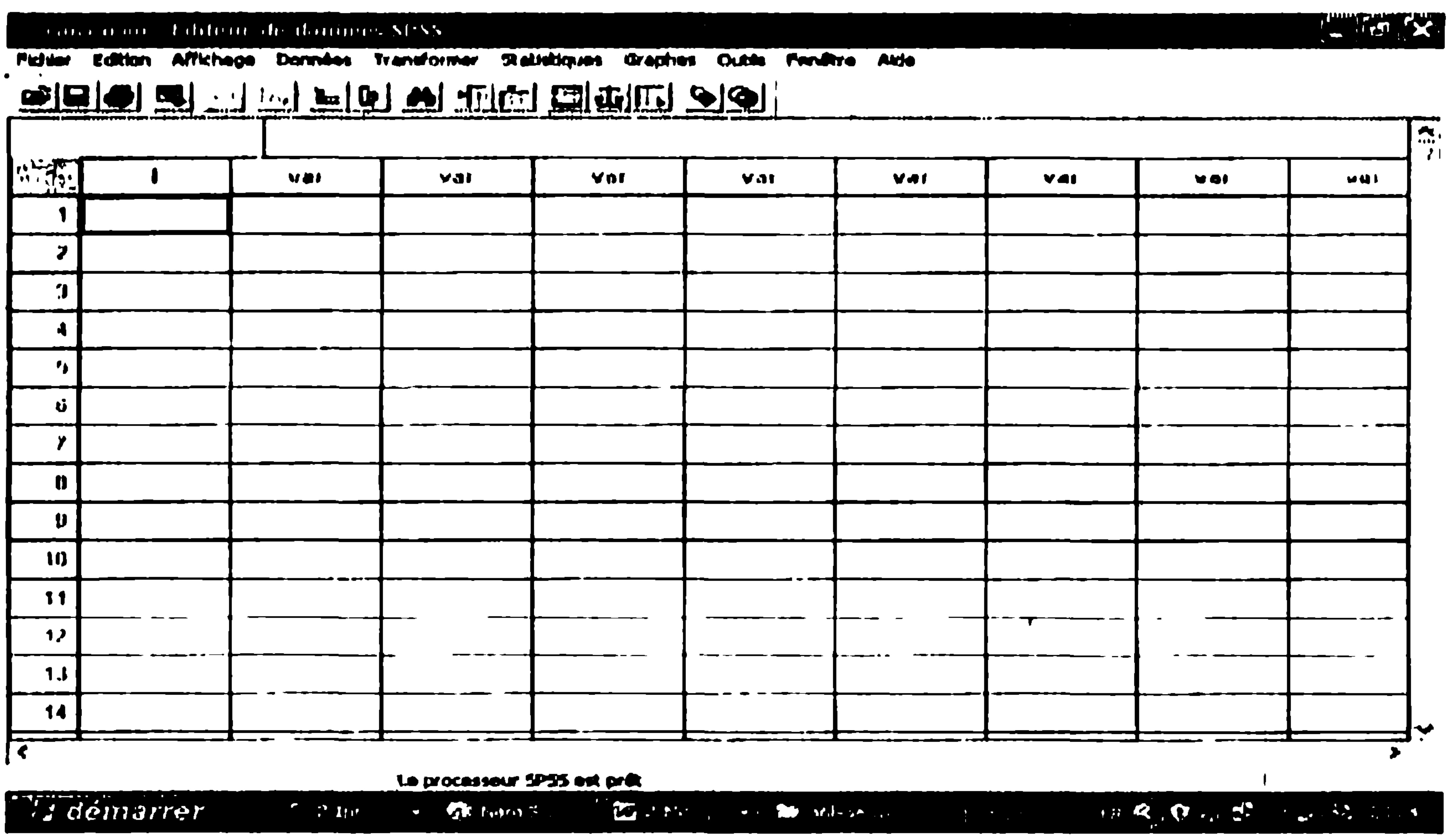
6. بالضغط نحصل على :



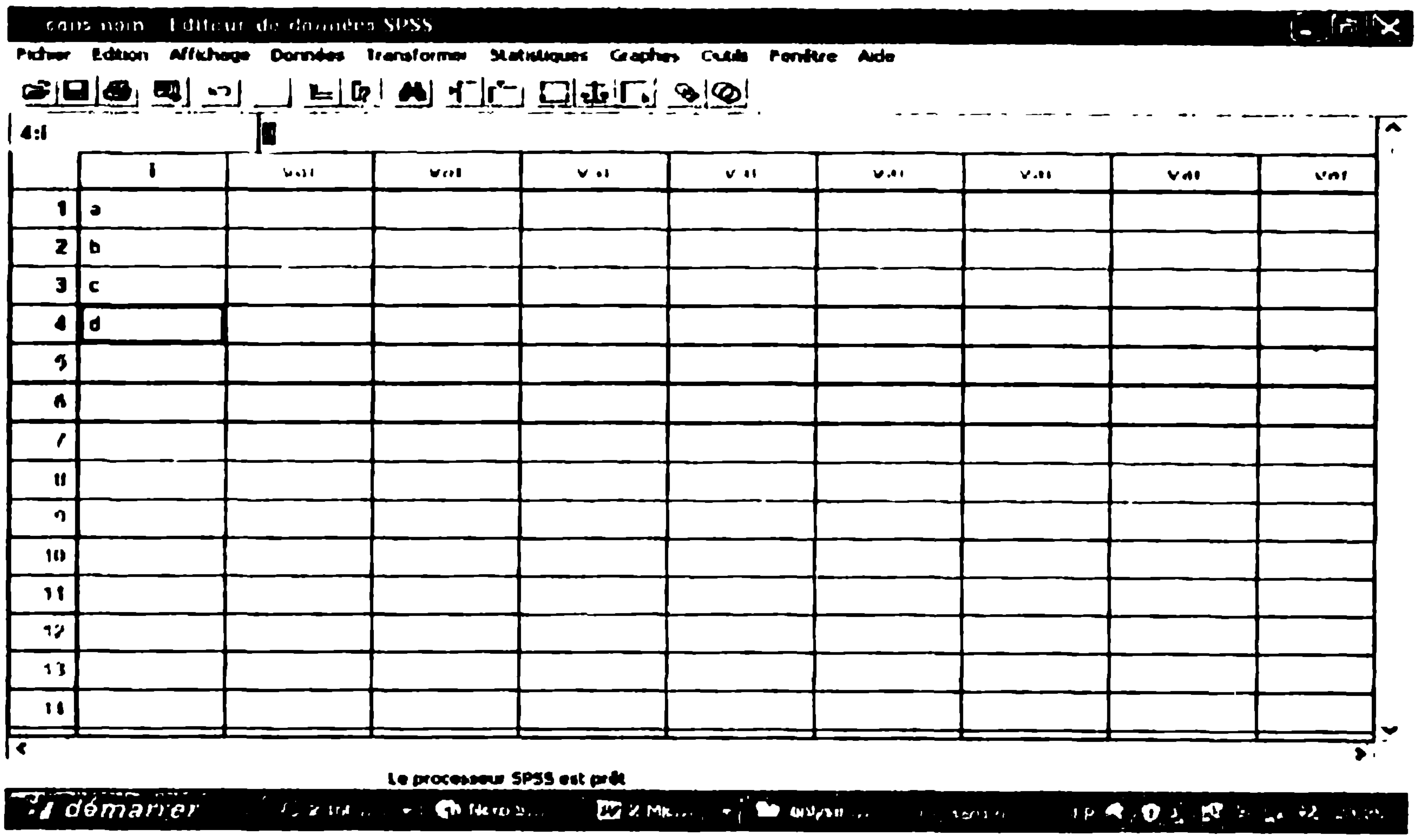
7. نقوم بتغيير VAR00001 بـ i ونقوم بالضغط على Type ونضغط على Chaine. كما يلي :



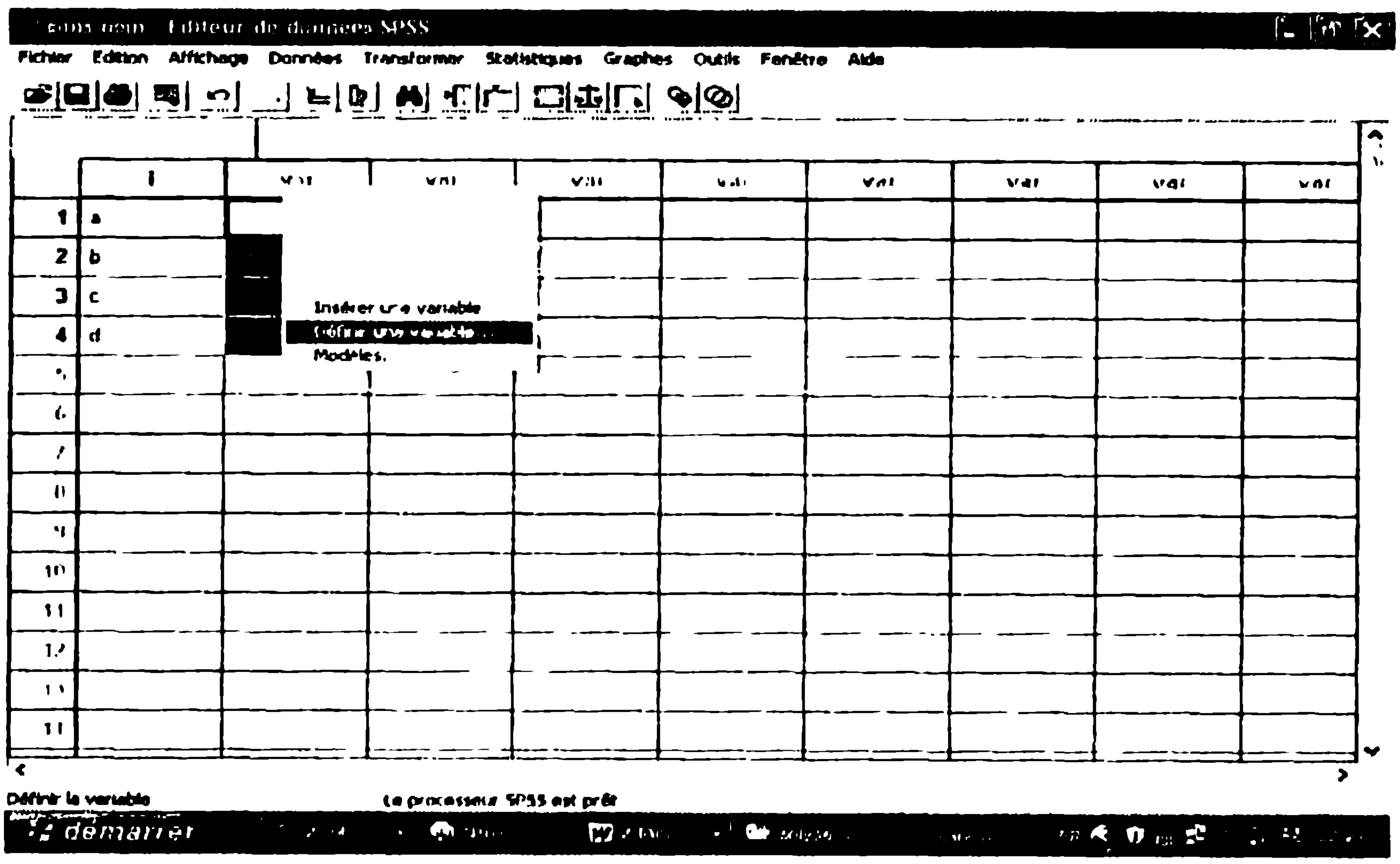
8. بالضغط على Poursuivre. ثم على OK نحصل على :



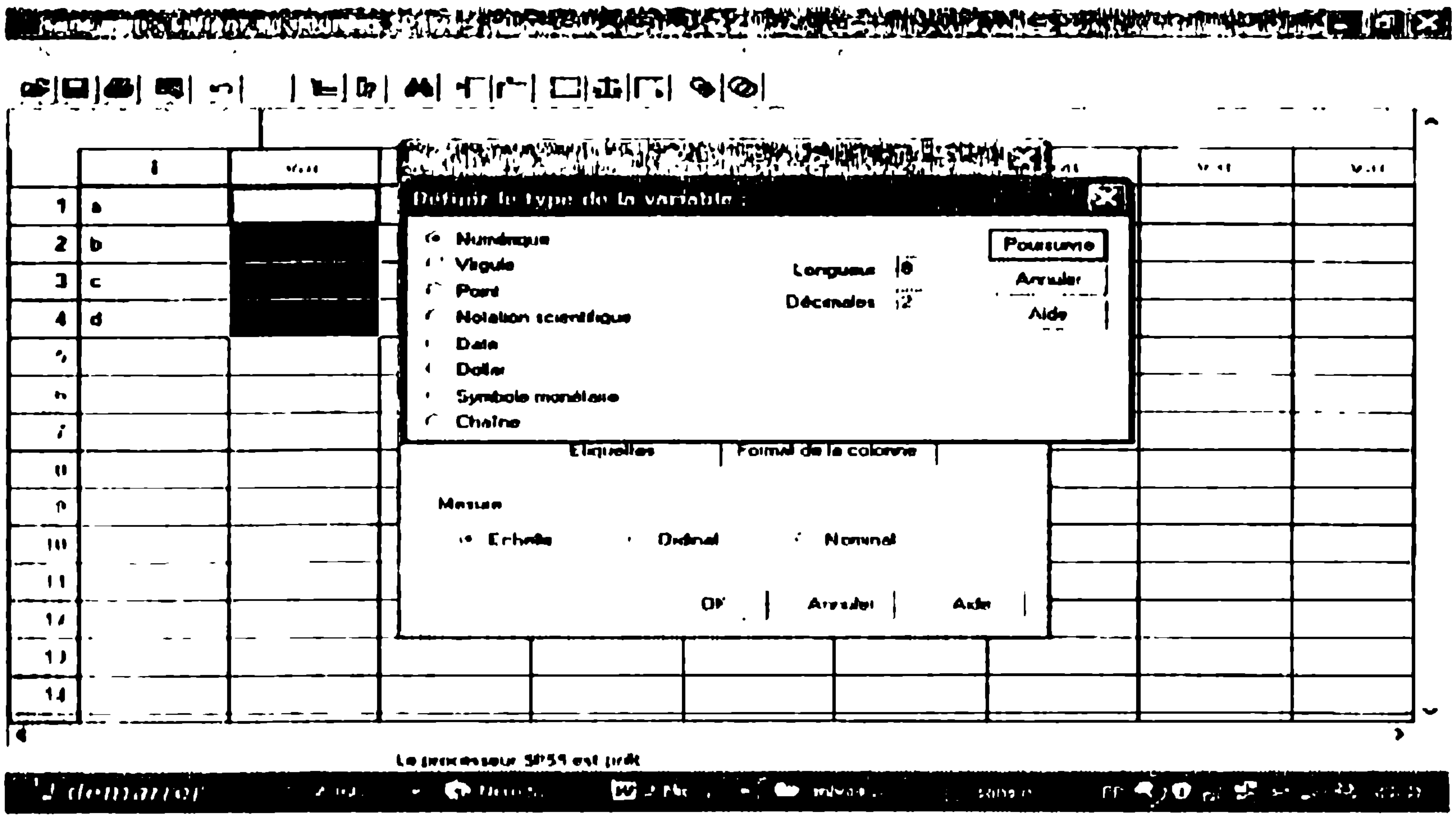
9. نقـوم بإدخـال الأفـراد (a,b,c,d)



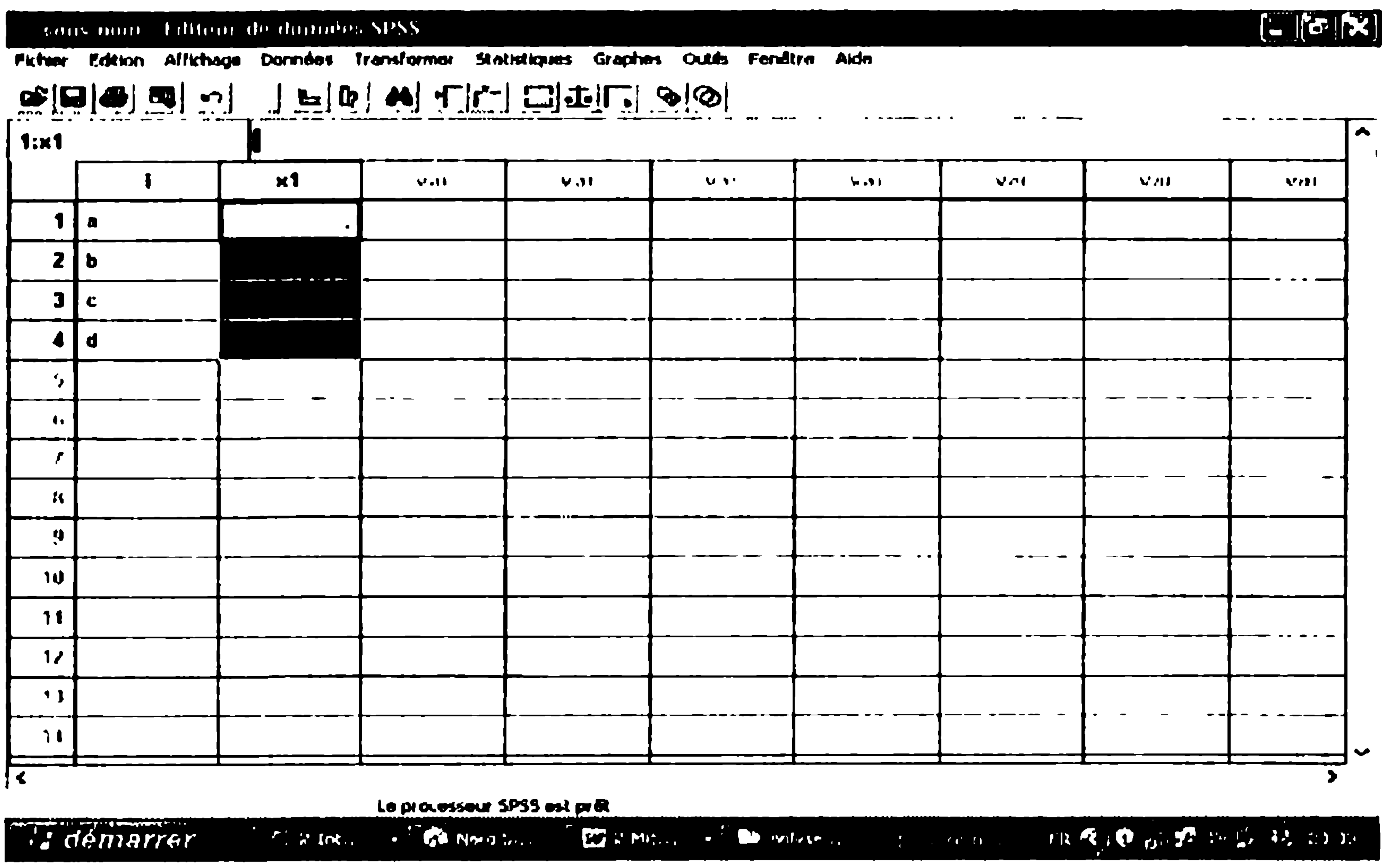
10. نذهب إلى العمود الثاني للشاشة ونقوم بتحديده، ونضغط على الزر الأيسر للمشيرة La sourie لادخال المتغيرة الأولى X1 كما يلي :



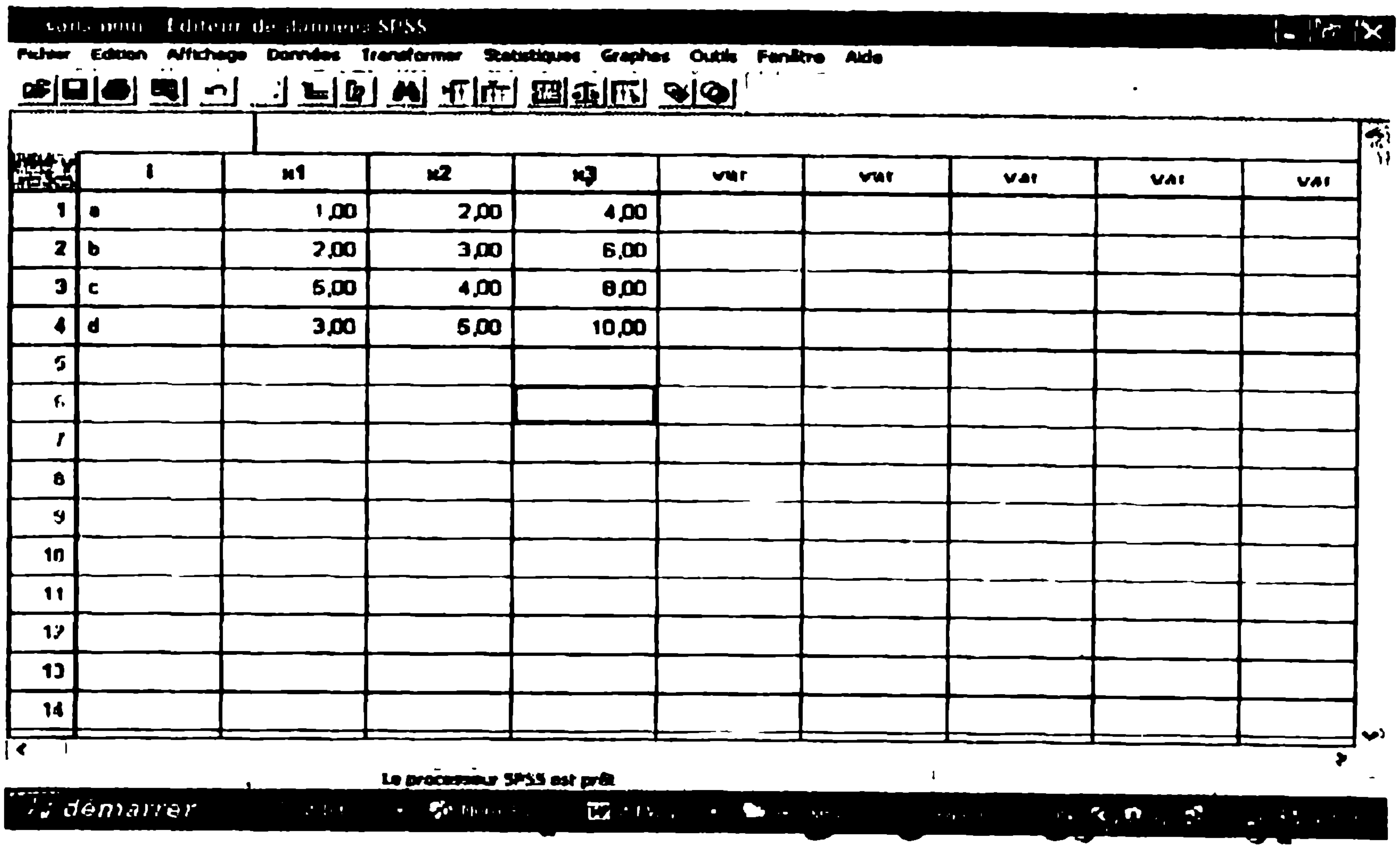
11. بالضغط على Définir une variable، نقوم بتسمية العمود الثاني X2، ونقوم بالضغط على Type، ولكن في هذه المرة نقوم بتحديد Numérique، ثم على Poursuivre، ثم على OK.



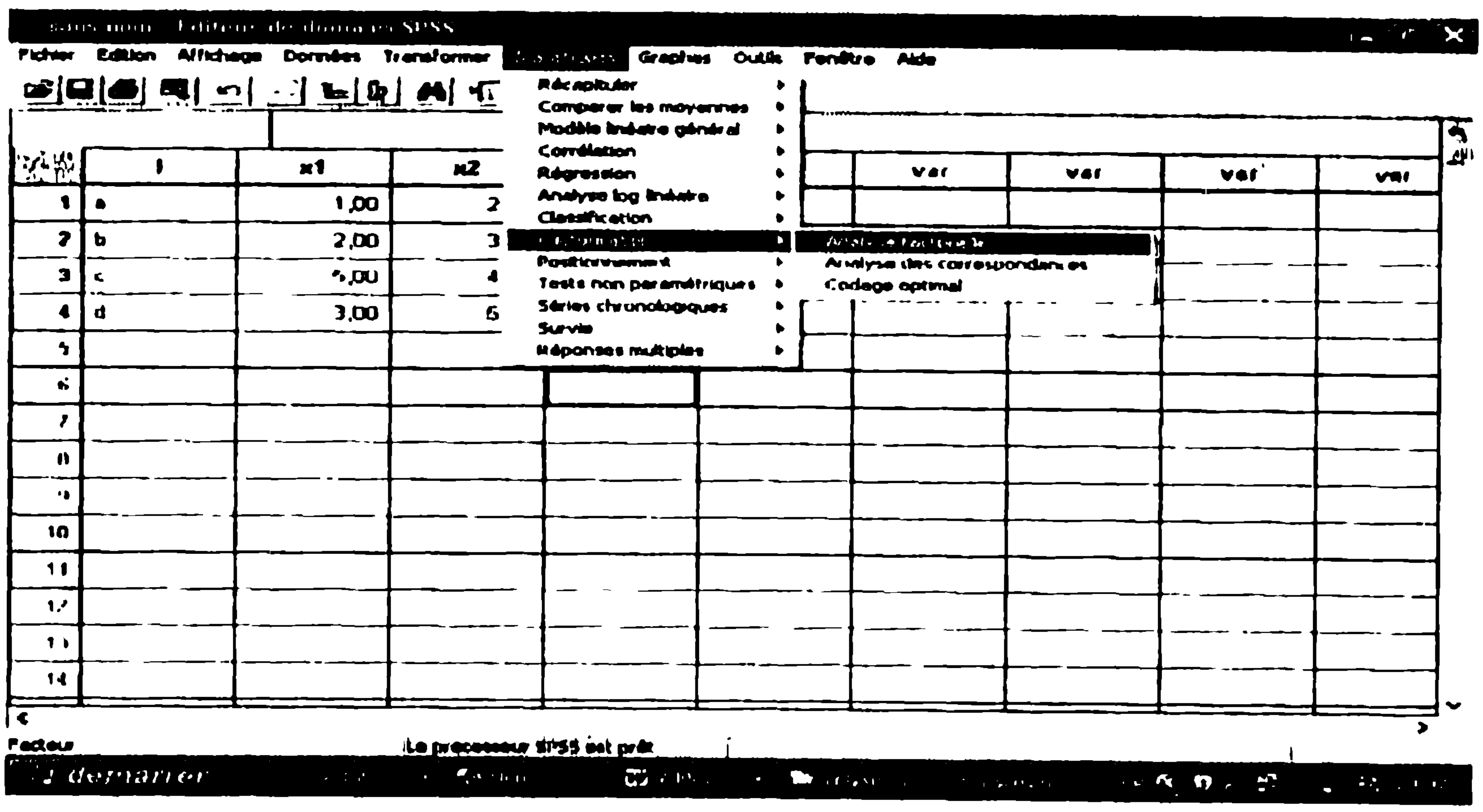
12. نحصل على :



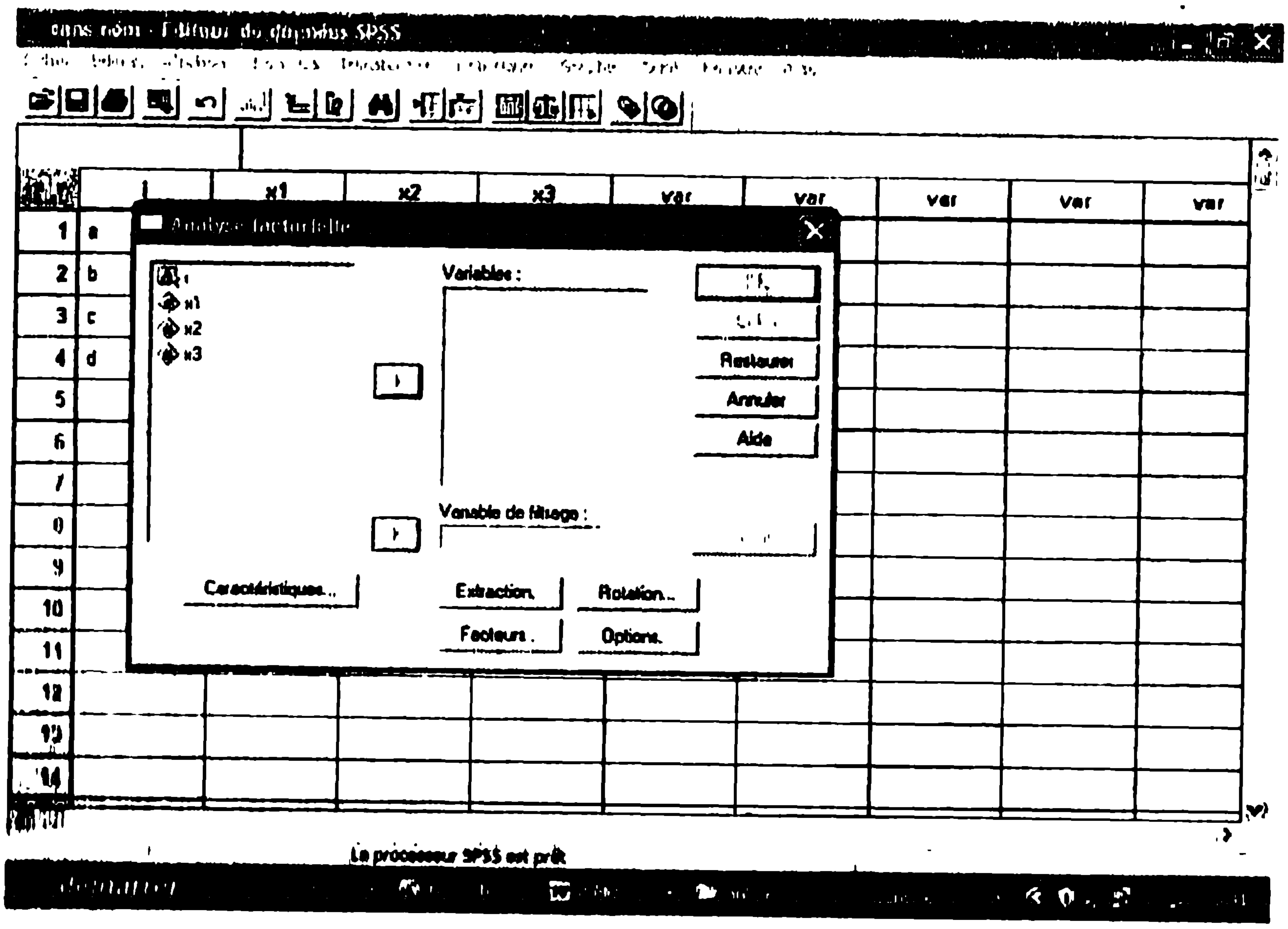
13. ثم نقوم بإدخال المتغيرة الأولى، ثم نعيد نفس المراحل بالنسبة للمتغيرات الأخرى نحصل على الشاشة التالية :



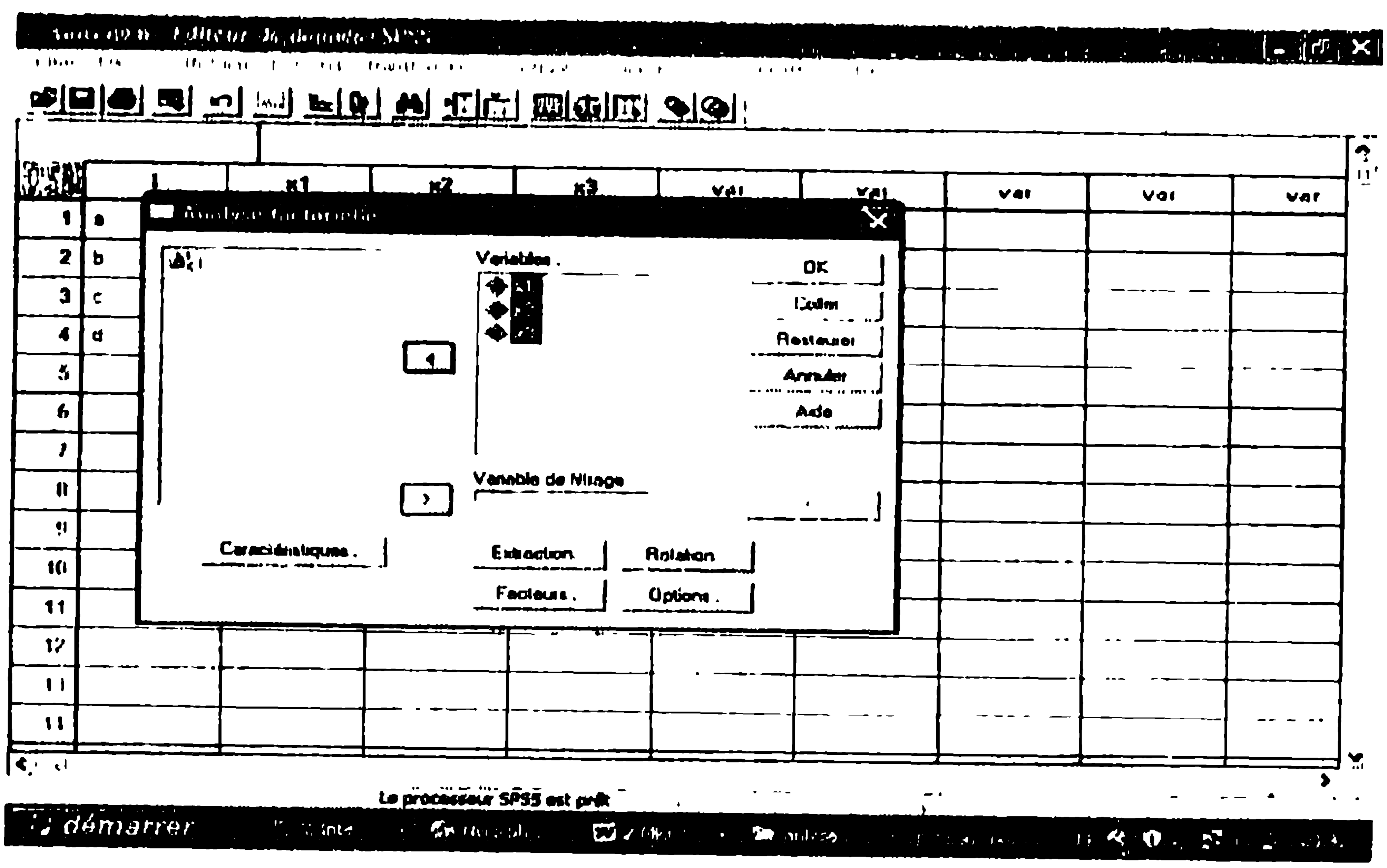
14. نذهب الآن إلى Statistiques، ثم Factorisation، ثم على Analyse factorielle كمايلي :



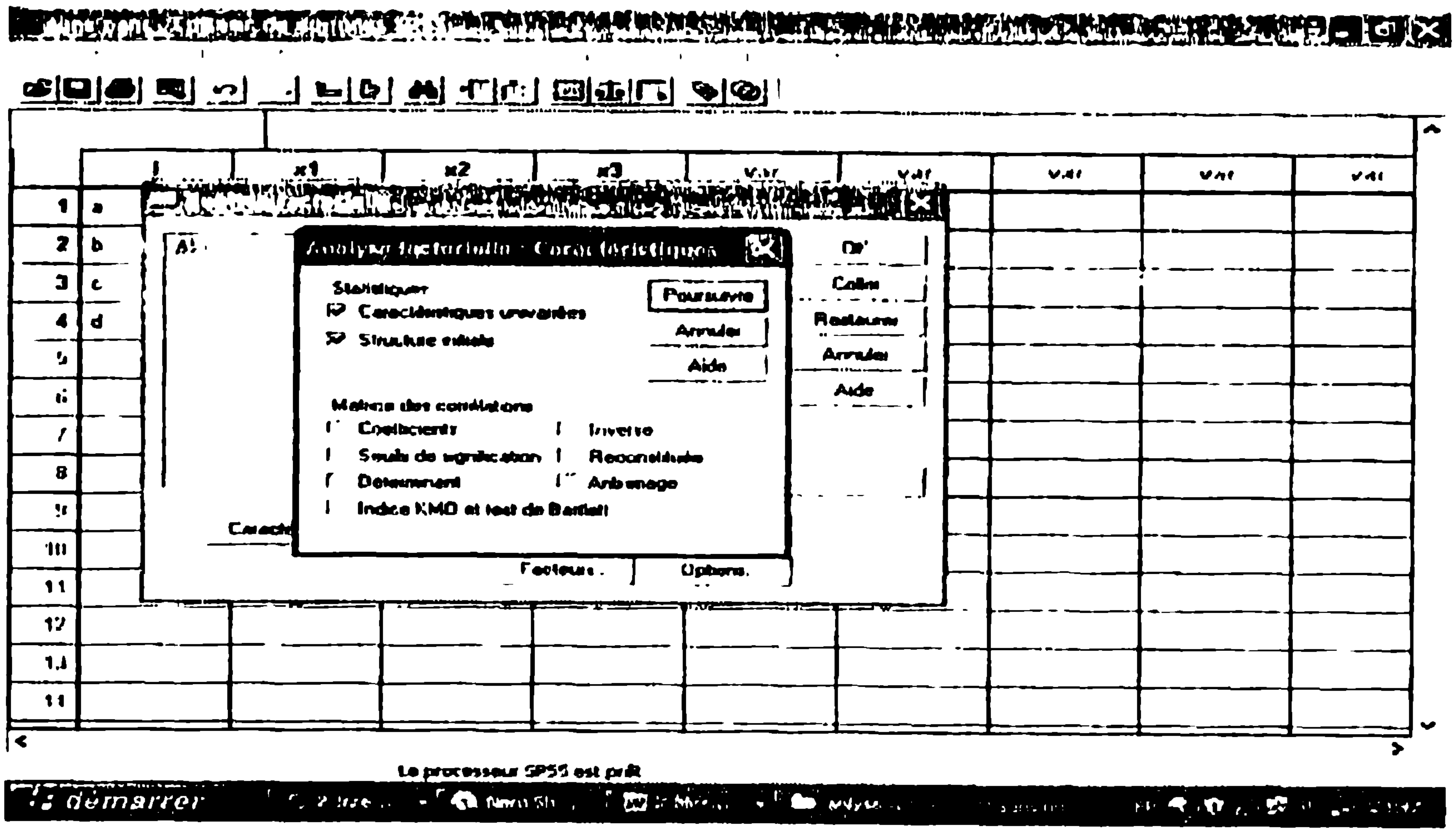
15. بالضغط على Analyse factorielle نحصل على :



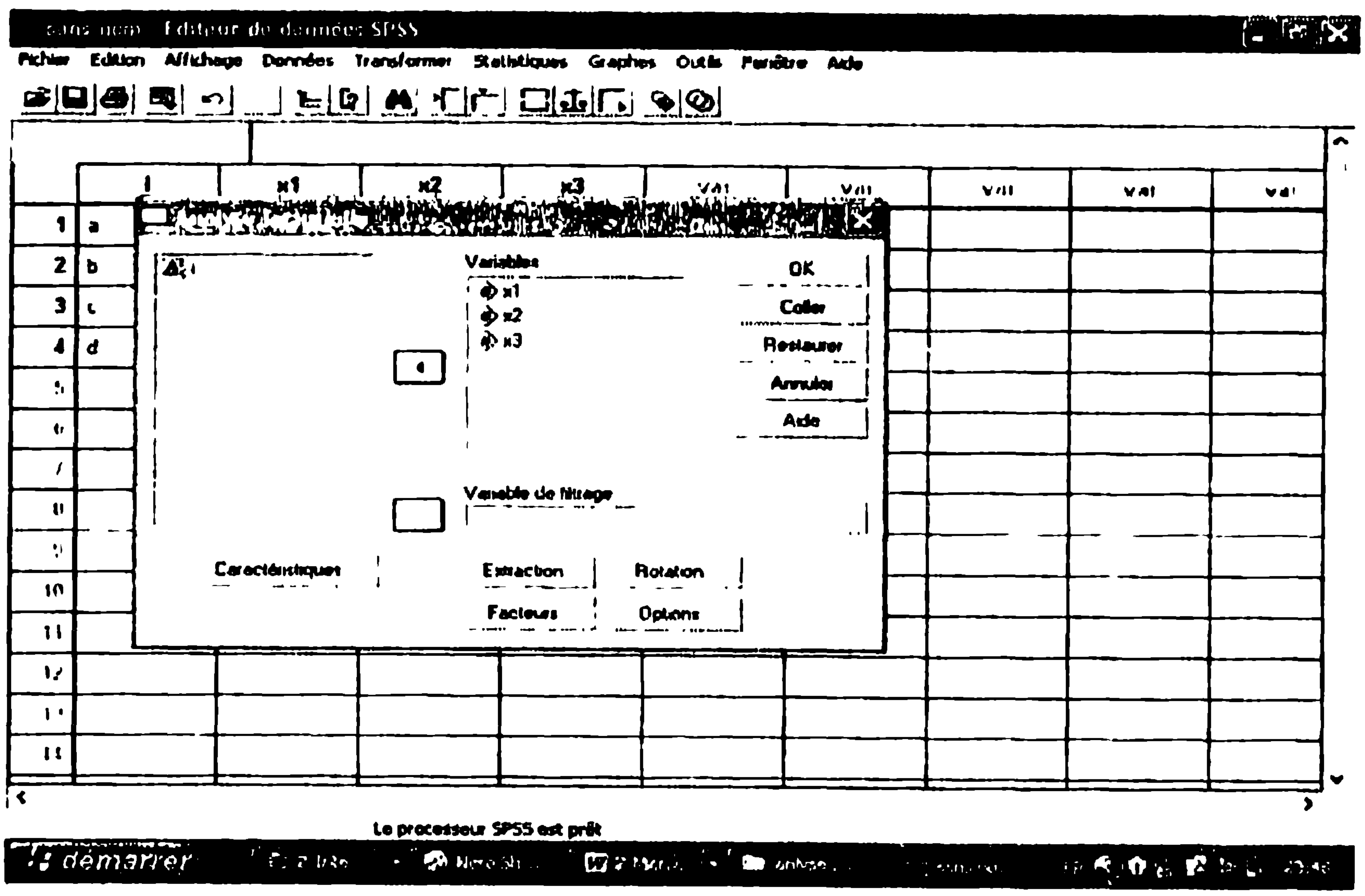
16. نقوم بإدخال المتغيرات X1، X2، X3 كما يلي :



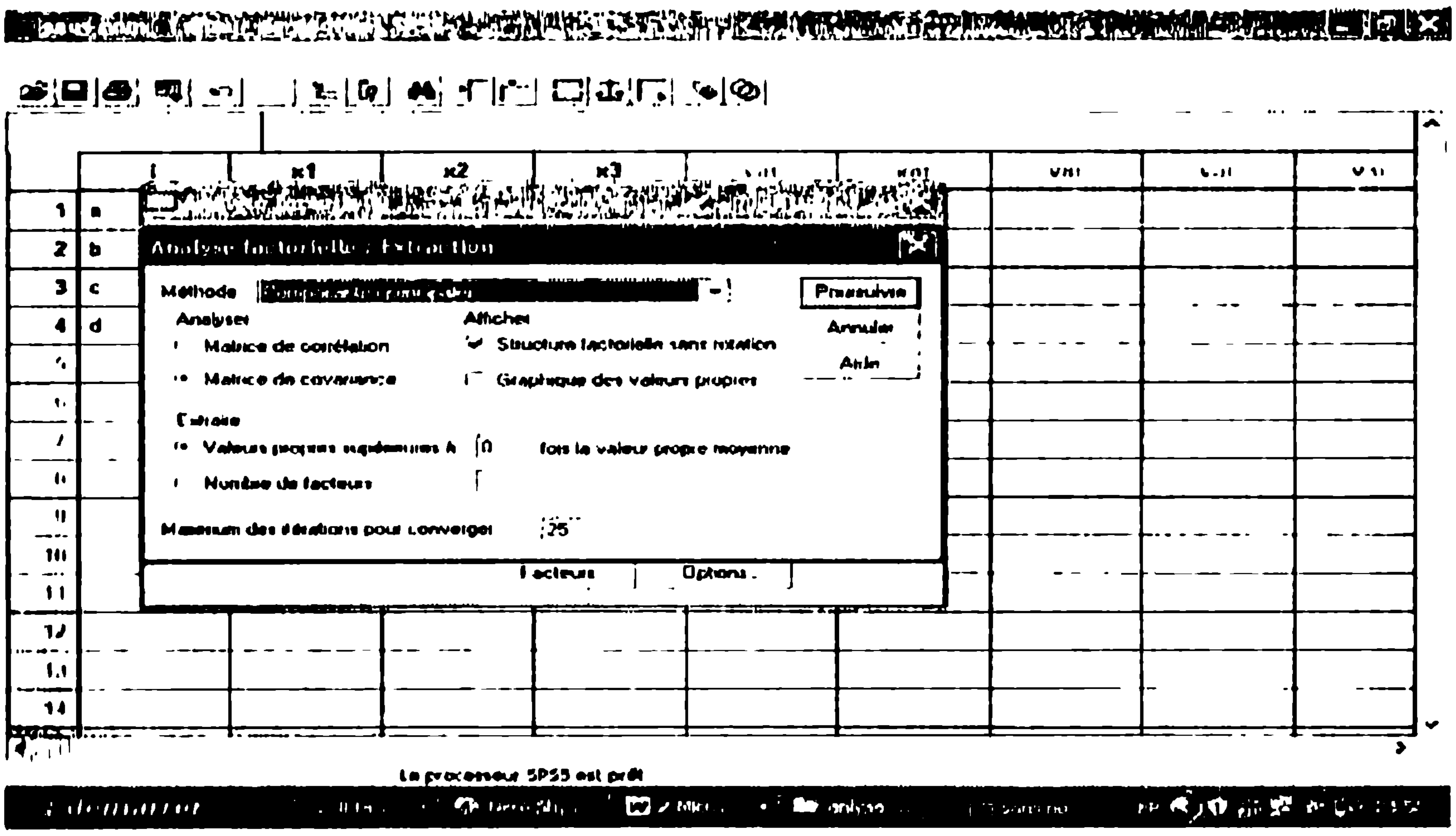
17. نذهب إلى Caractéristiques ونضغط عليها ثم نقوم بتحديد كل من Caractéristiques Uni variées، وعلى Structure initiale كما يلي :



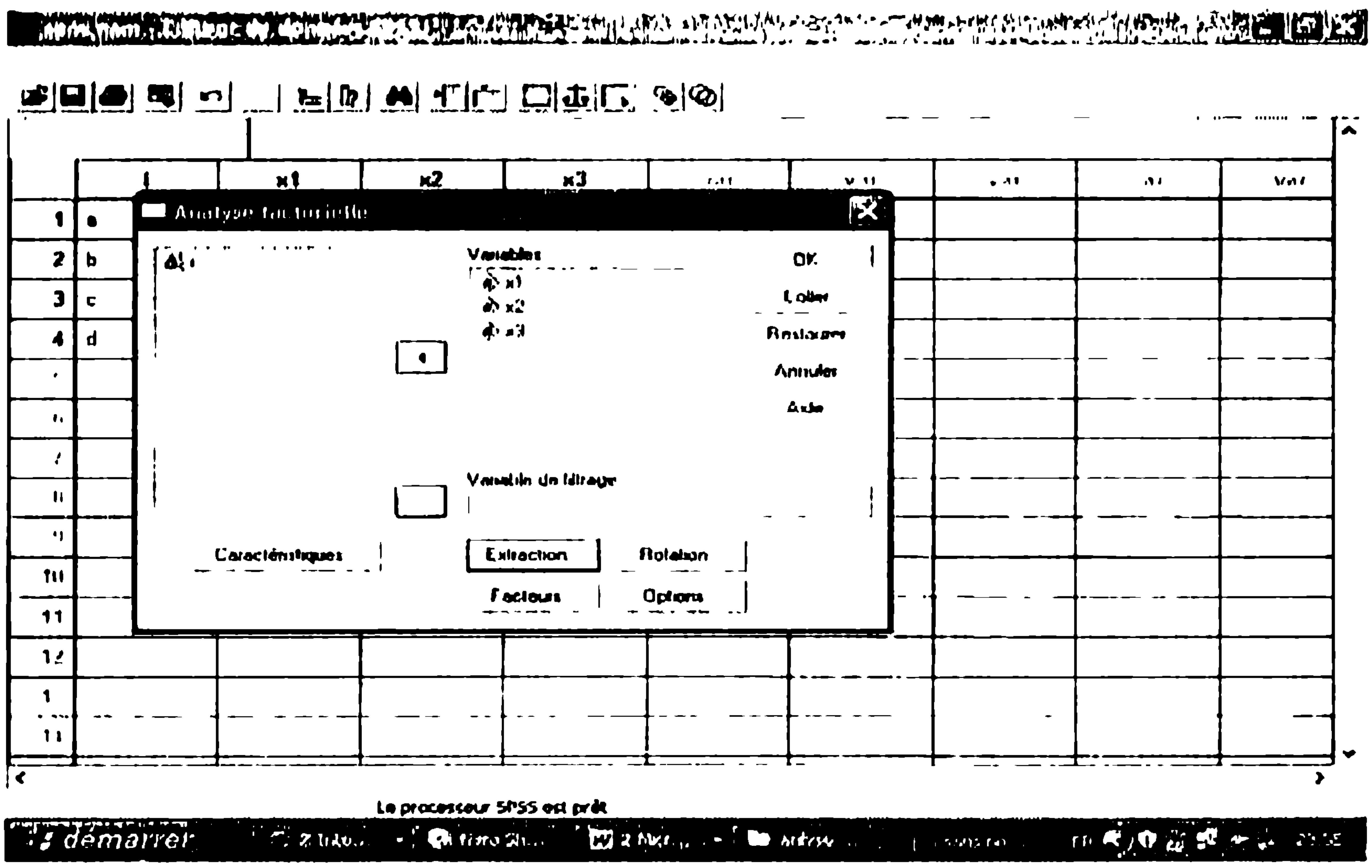
18. نقوم بضغط على Poursuivre نحصل على :



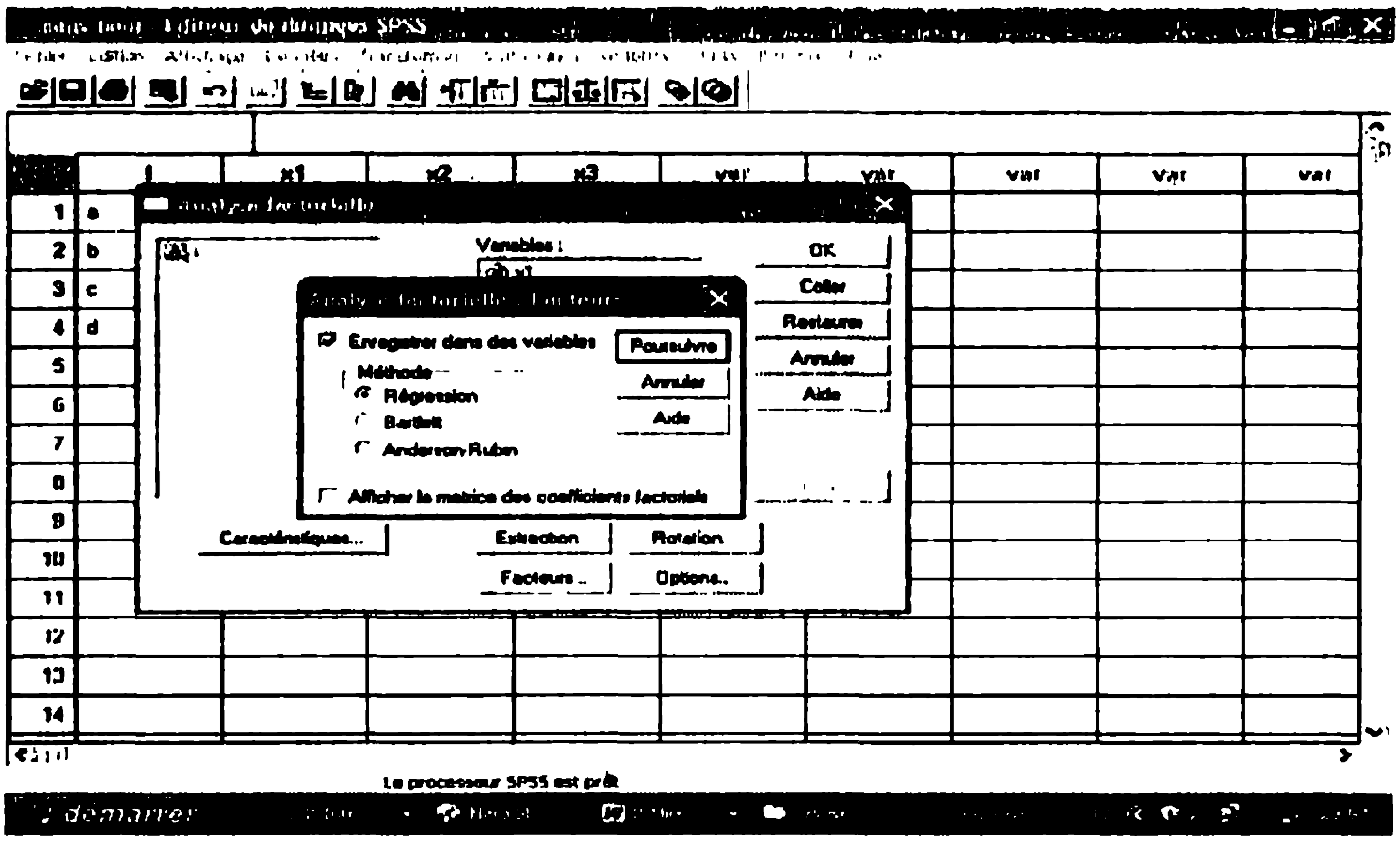
19. نذهب إلى Extraction ونقوم بتحديد Méthode - composante principale - ، ثم Matrice de covariance، ثم Valeurs Propre supérieur à 0، كما يلي :



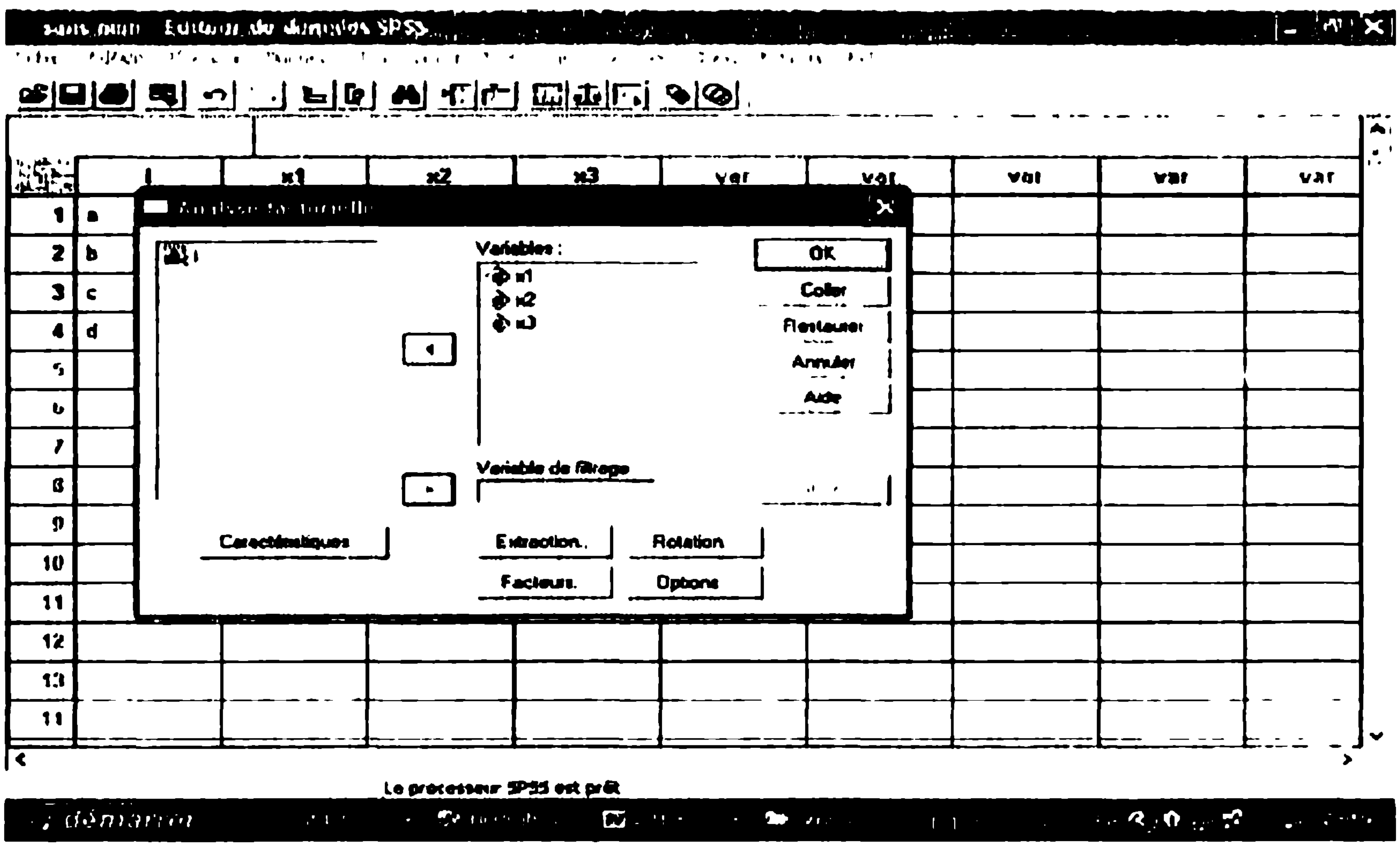
20. نضغط على Poursuivre



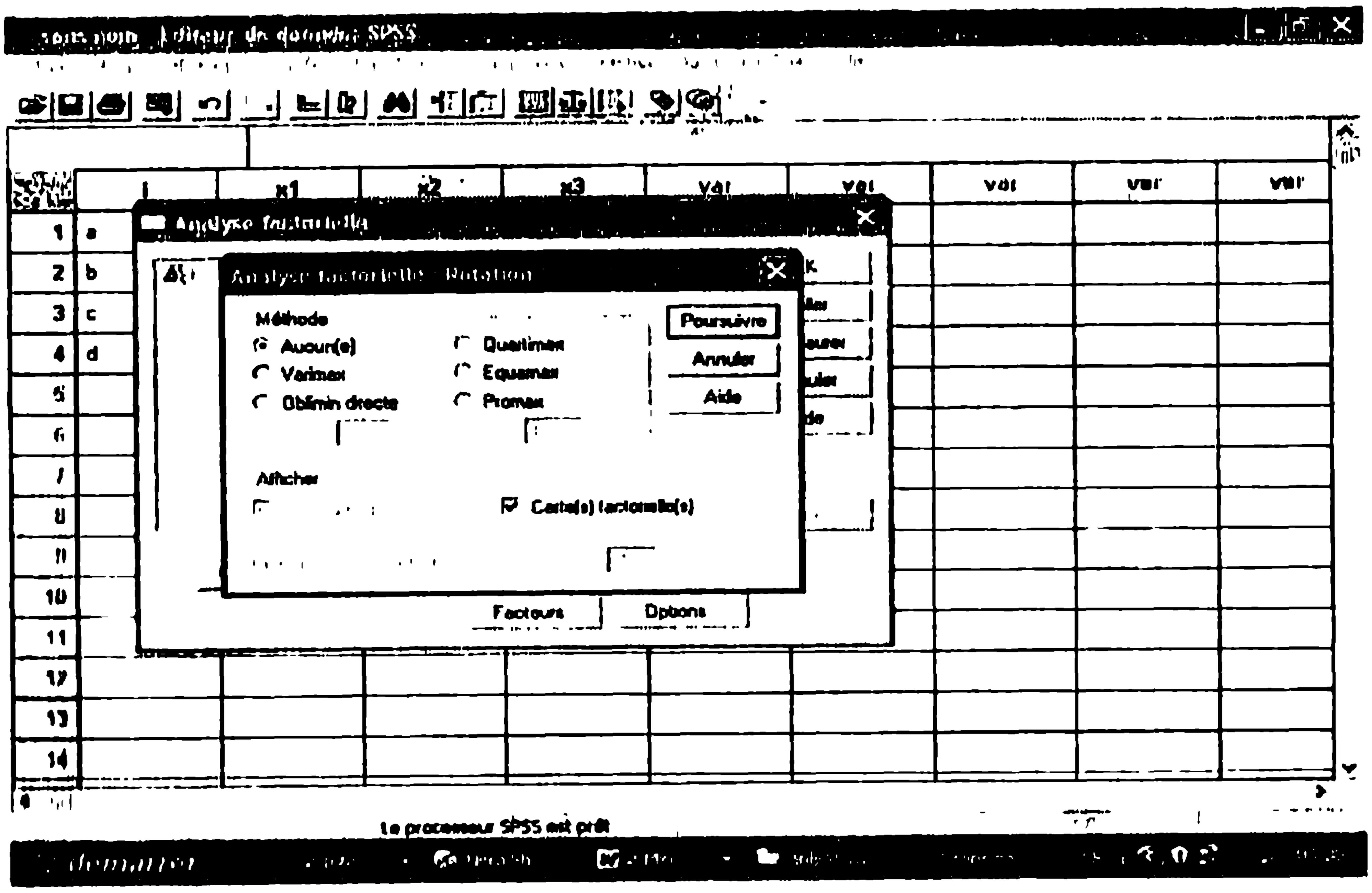
21. نذهب إلى Facteurs ، ونقوم بالضغط على Enregistres dans des variables كما يلي :



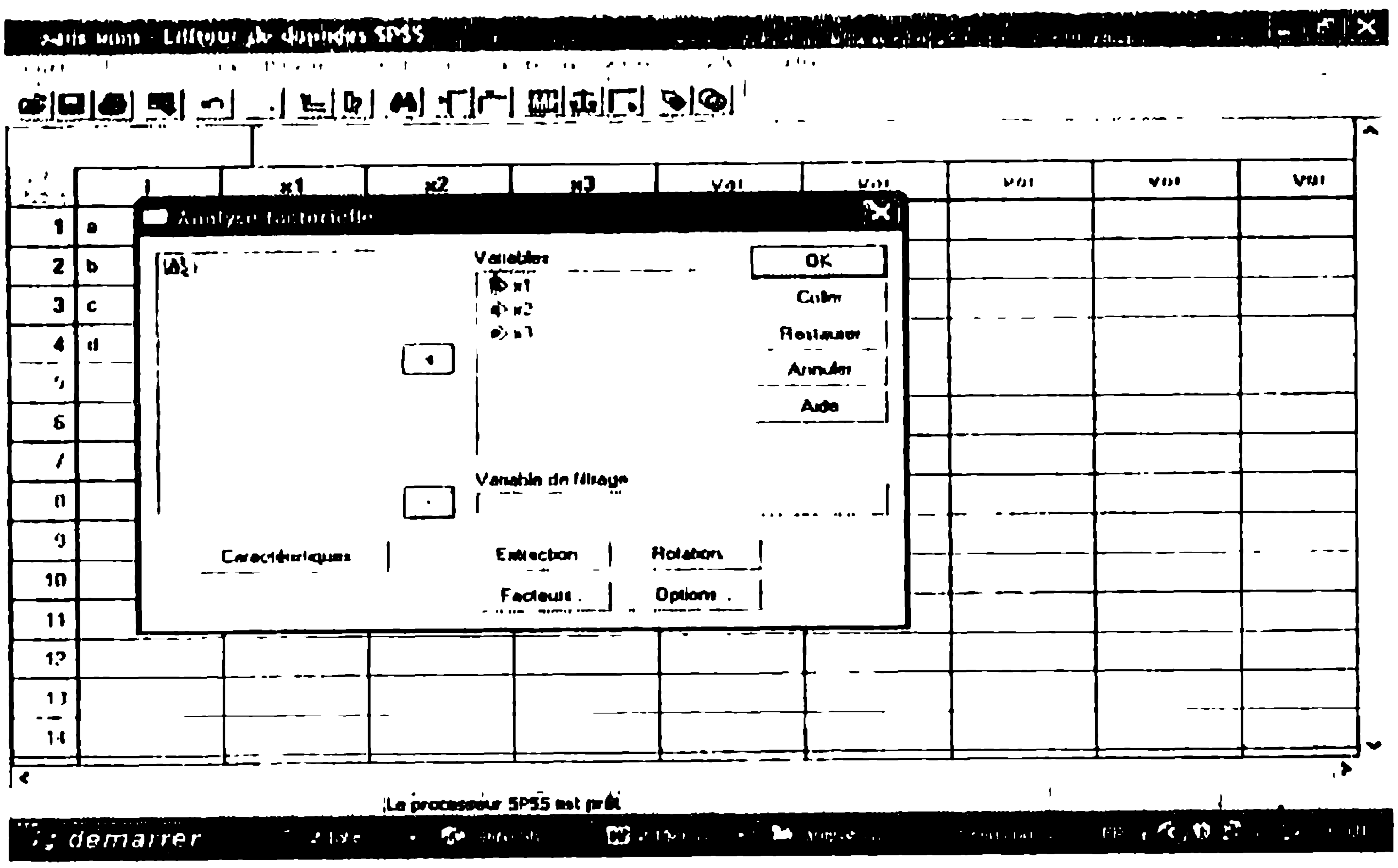
22. نضغط على Poursuivre ، نحصل على :



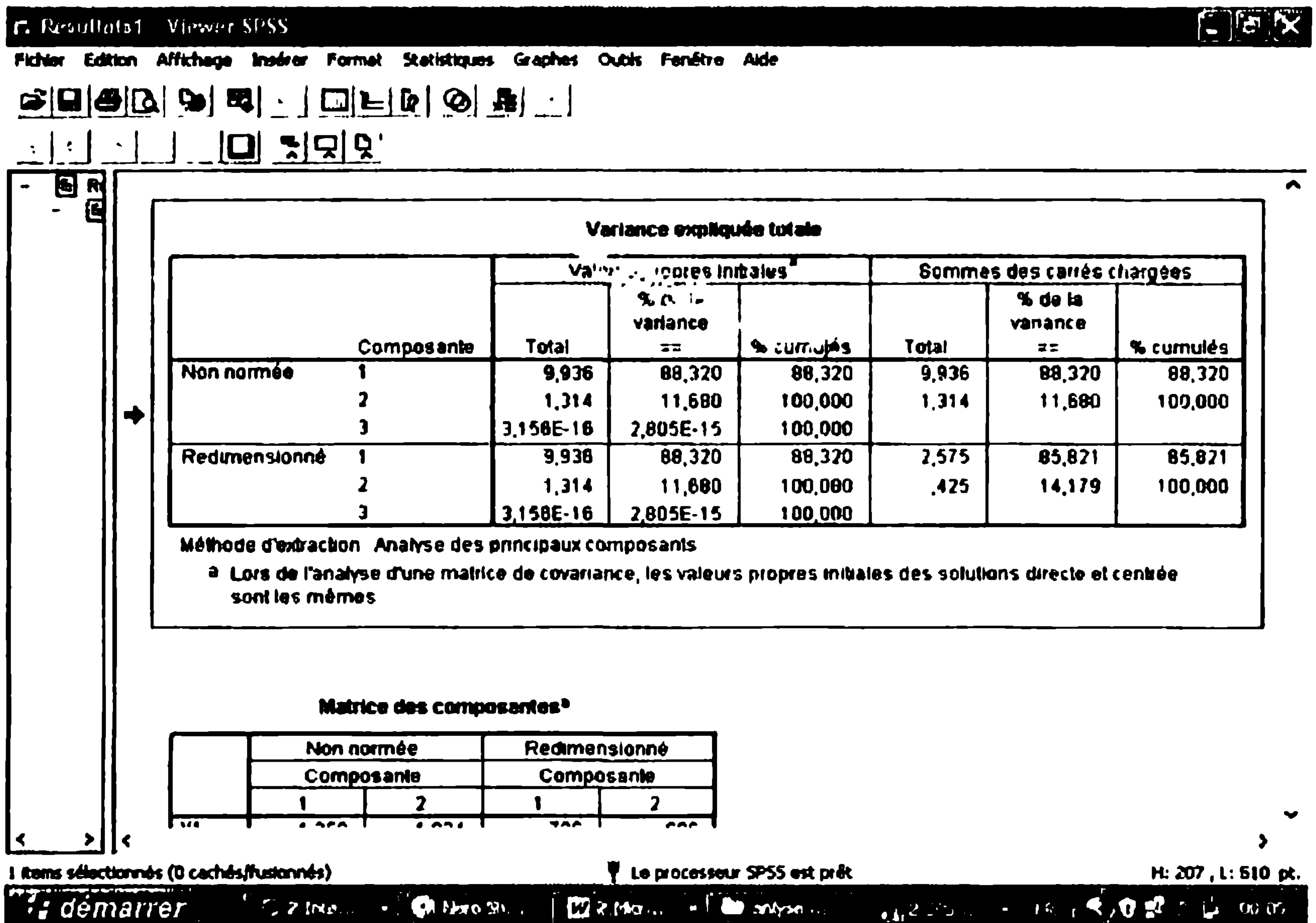
23. نذهب إلى Rotation، نضغط على Aucune وعلى Carte(s) factorielle(s) كما يلي :



24. نحصل على :



25. نقوم بالضغط على OK، نحصل على الشاشة التالية والتي تحتوي على القيم الذاتية وتمثيل المتغيرات :



Variance expliquée totale

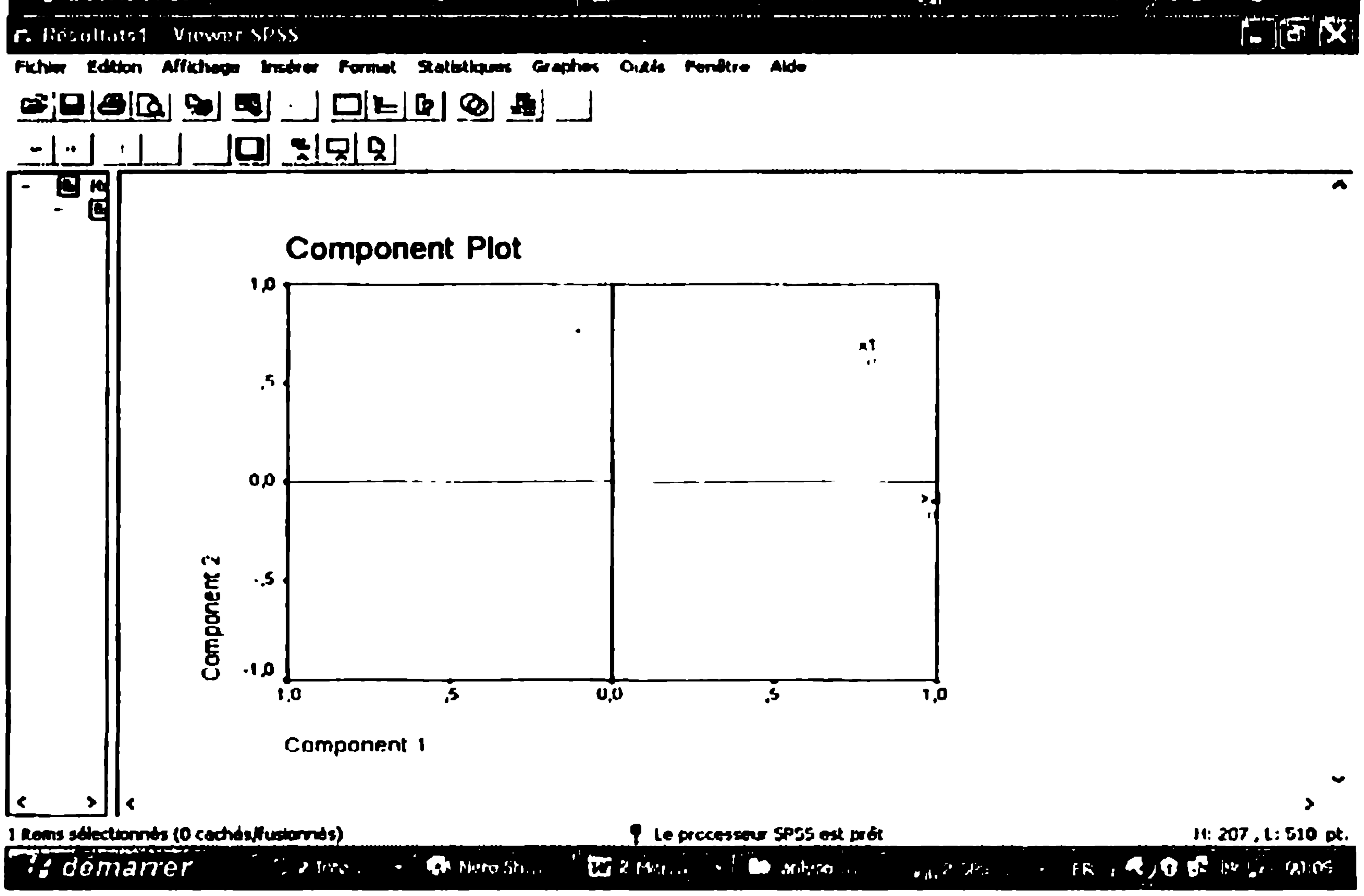
	Composante	Val[illegible]propres initiales[a]			Sommes des carrés chargées		
		Total	% de la variance	% cumulés	Total	% de la variance	% cumulés
Non normée	1	9,936	88,320	88,320	9,936	88,320	88,320
	2	1,314	11,680	100,000	1,314	11,680	100,000
	3	3,156E-16	2,805E-15	100,000			
Redimensionné	1	9,936	88,320	88,320	2,575	85,821	85,821
	2	1,314	11,680	100,000	,425	14,179	100,000
	3	3,156E-16	2,805E-15	100,000			

Méthode d'extraction Analyse des principaux composants

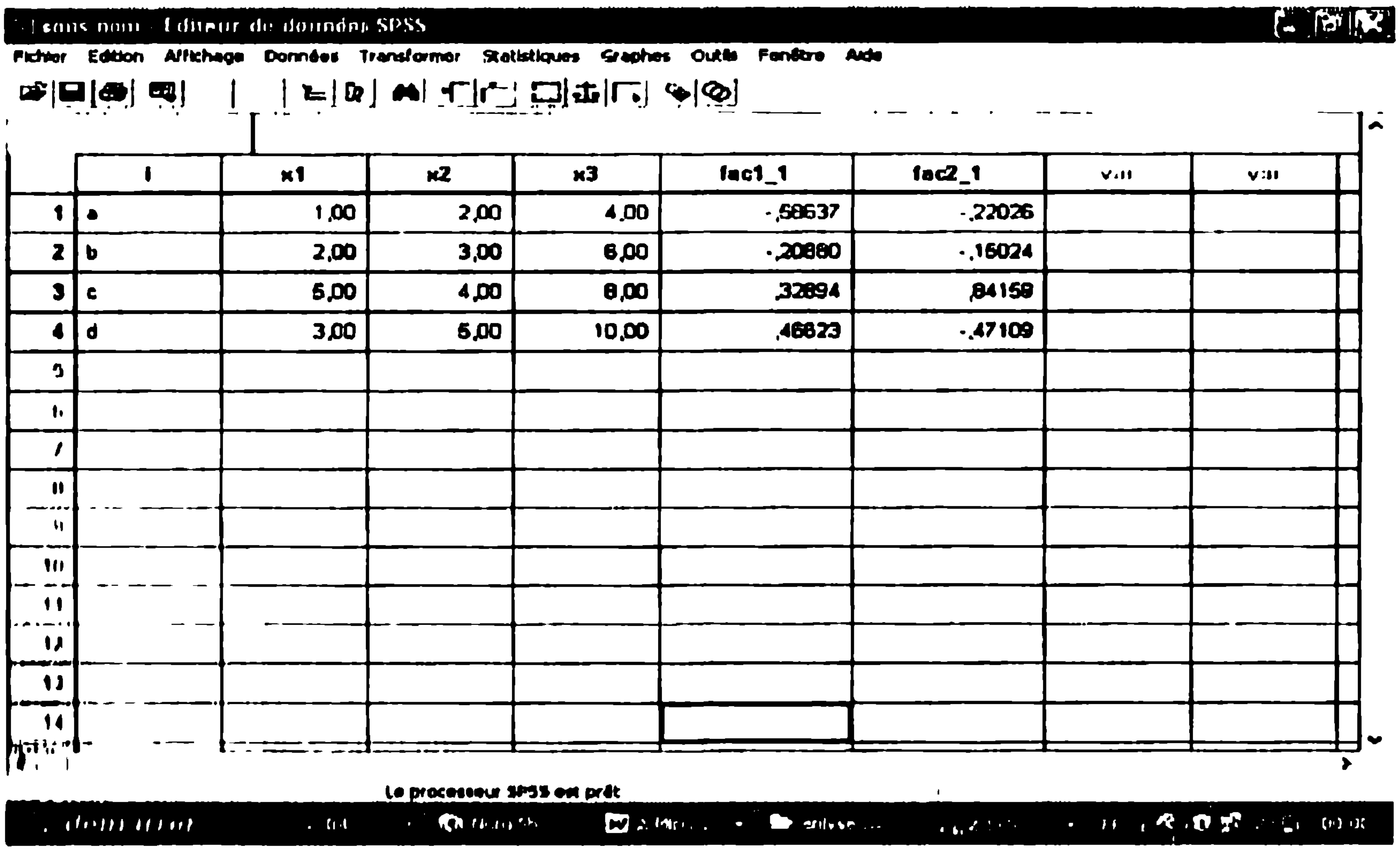
a. Lors de l'analyse d'une matrice de covariance, les valeurs propres initiales des solutions directe et centrée sont les mêmes

Matrice des composantes[a]

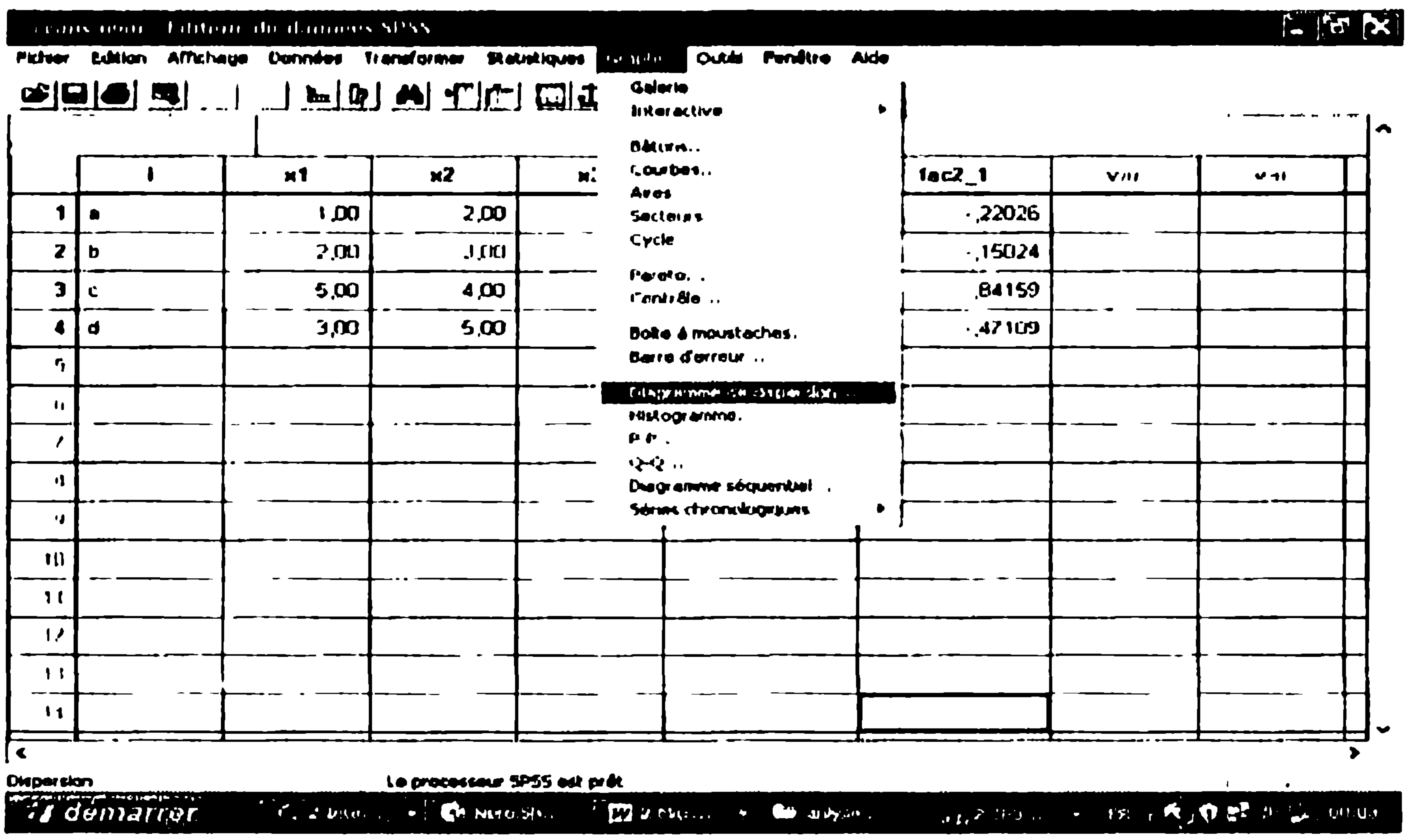
	Non normée		Redimensionné	
	Composante		Composante	
	1	2	1	2



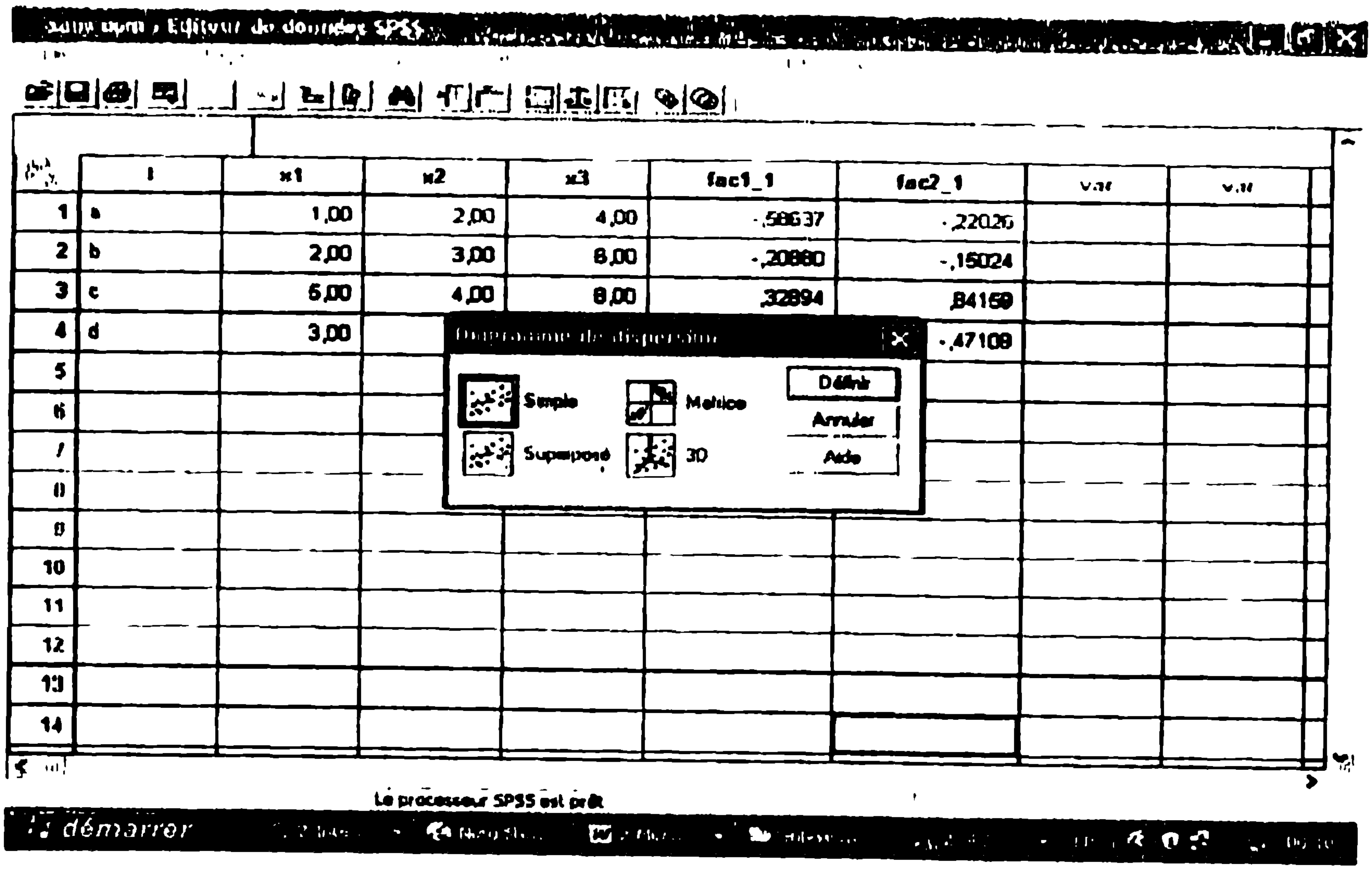
26. بالنسبة للمركبات الرئيسية للأفراد نجدها في الشاشة التي أدخلنا فيها المعطيات كما يلي :



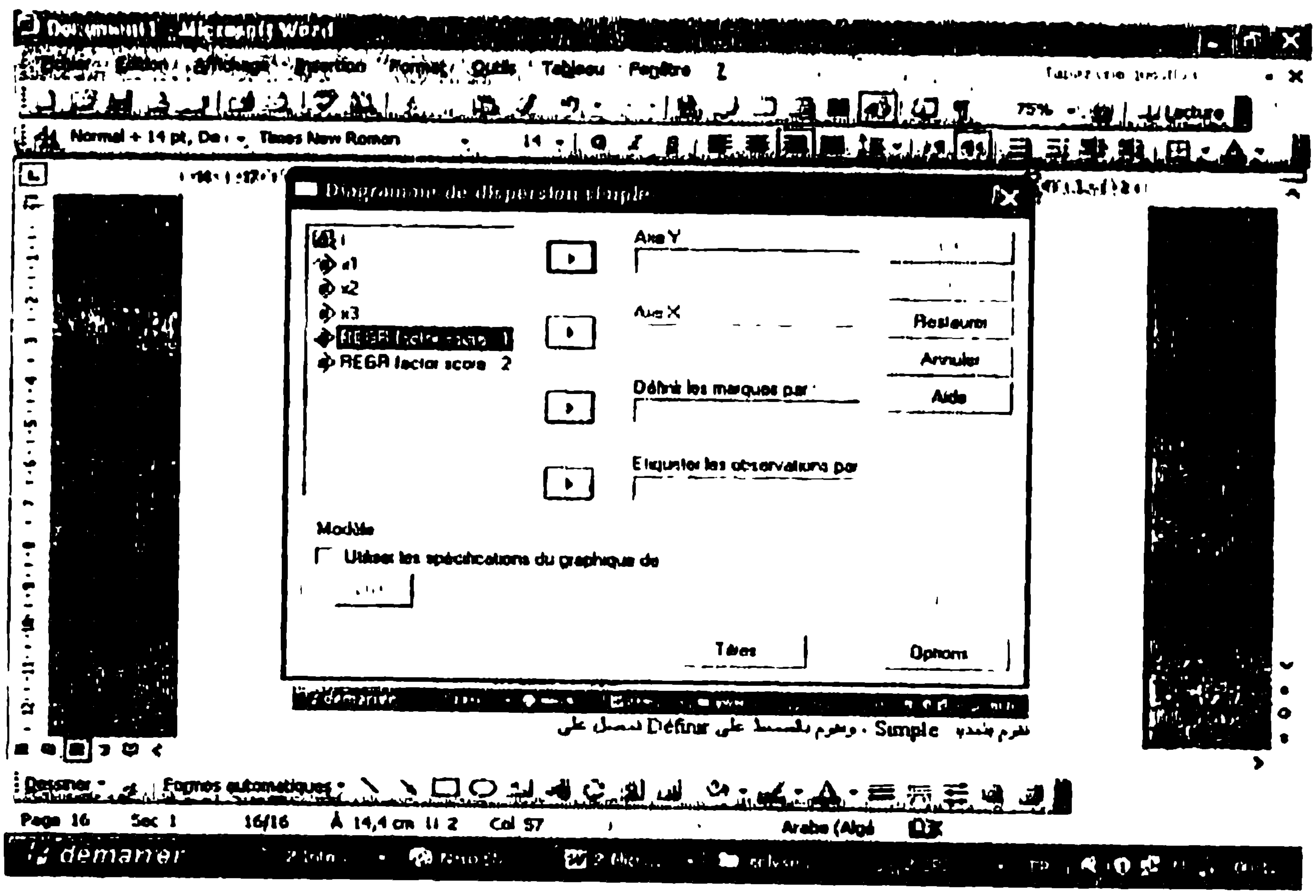
27. من أجل تمثيل الأفراد نذهب إلى Graphiques، Digramme de dispersion كما يلي :



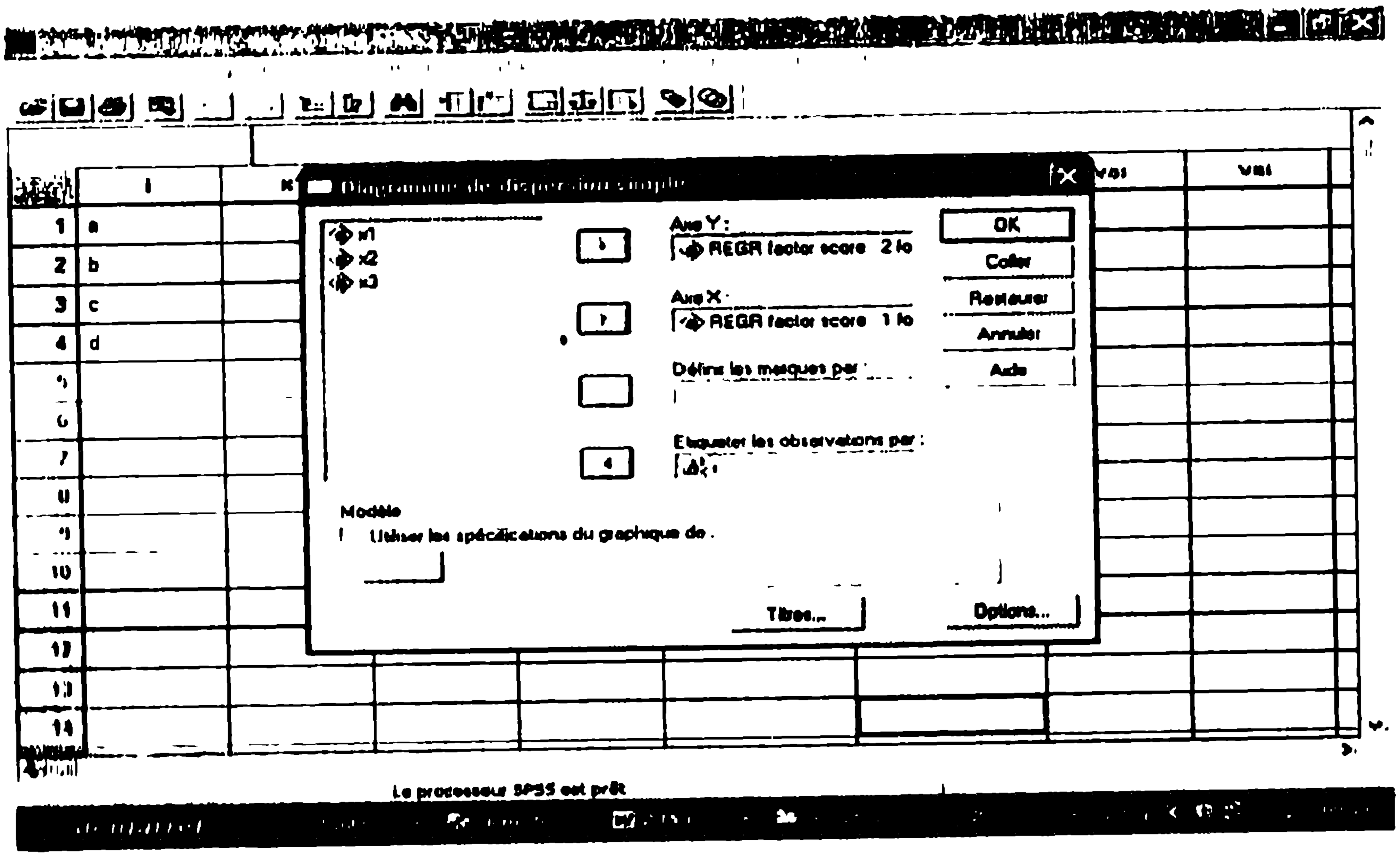
28. بالضغط عليها نحصل على :



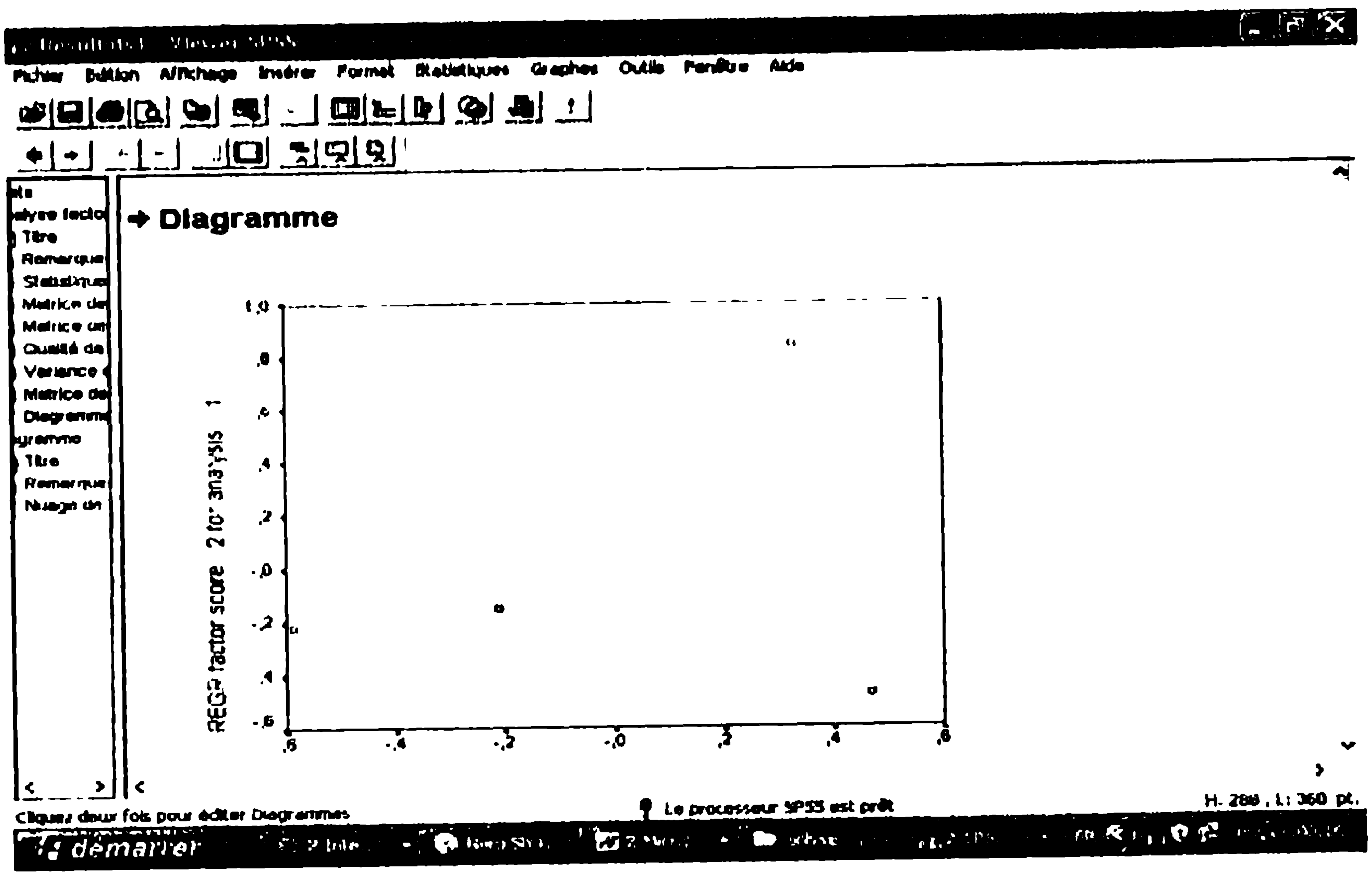
29. نقوم بتحديد Simple ، ونقوم بالضغط على Définir نحصل على :



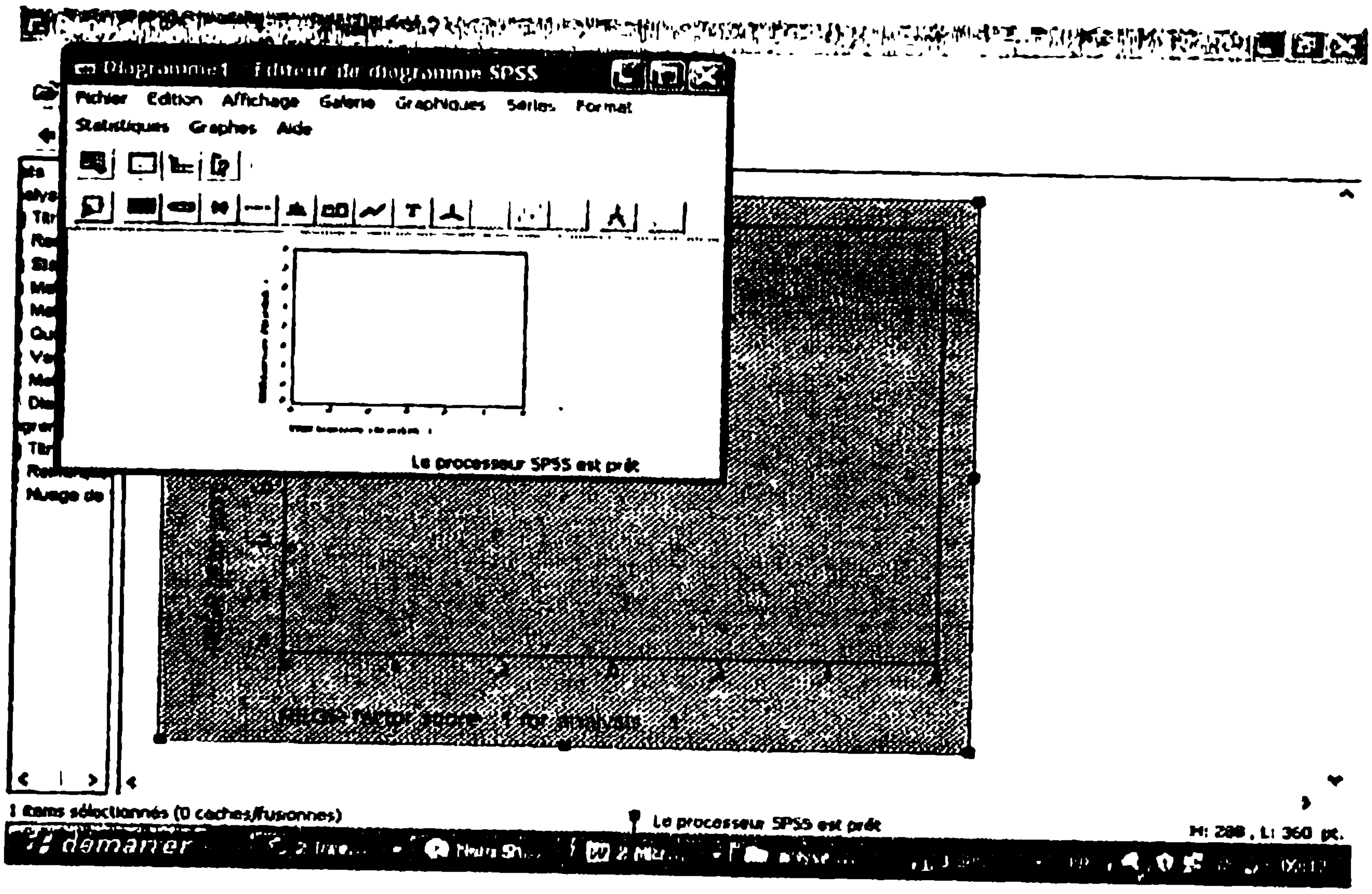
30. نقوم بتحديد 1 reger في X و2 reger في Y، وA في Etiqueter les observations par كما يلي :



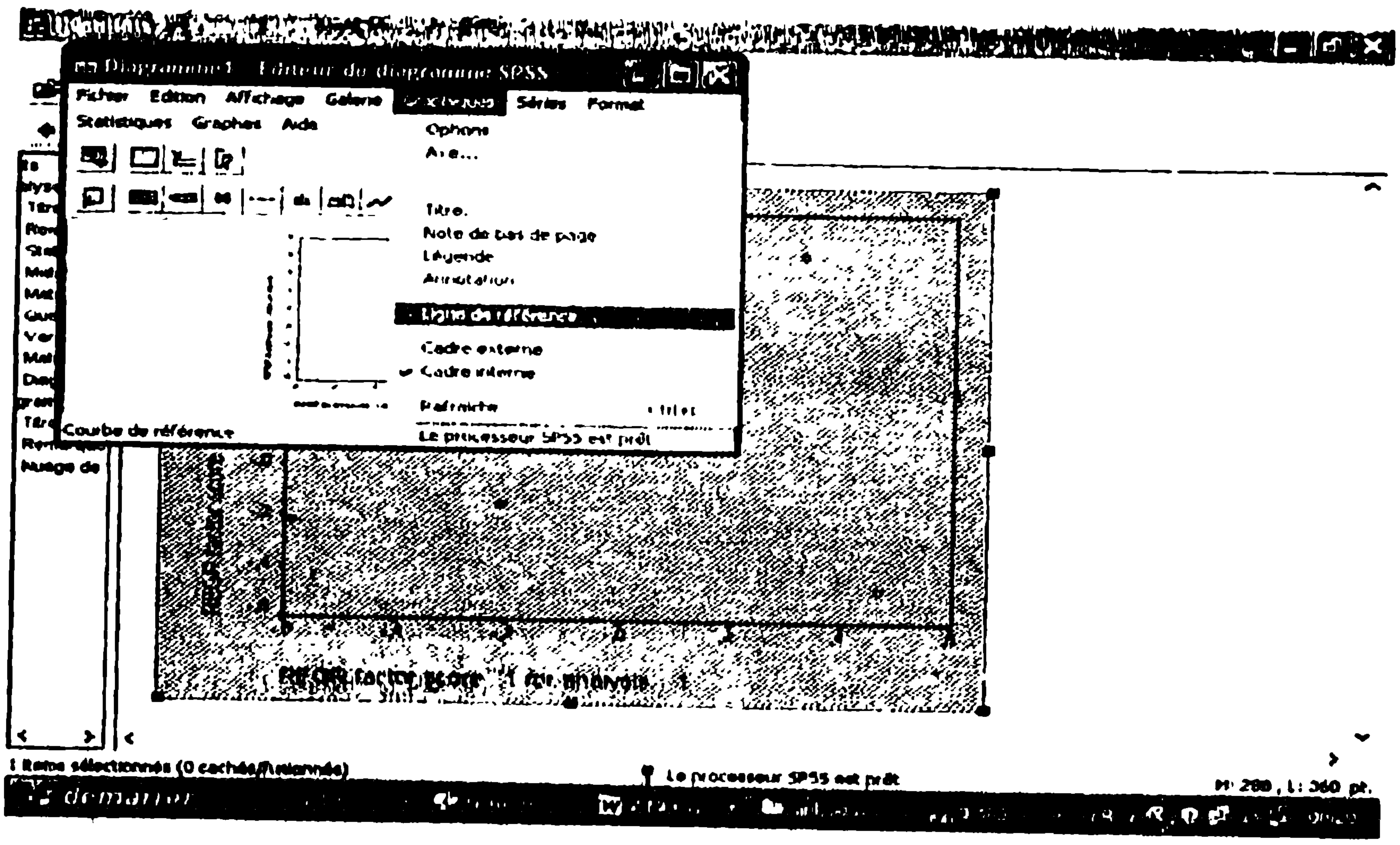
30. بالضغط على OK نحصل على :



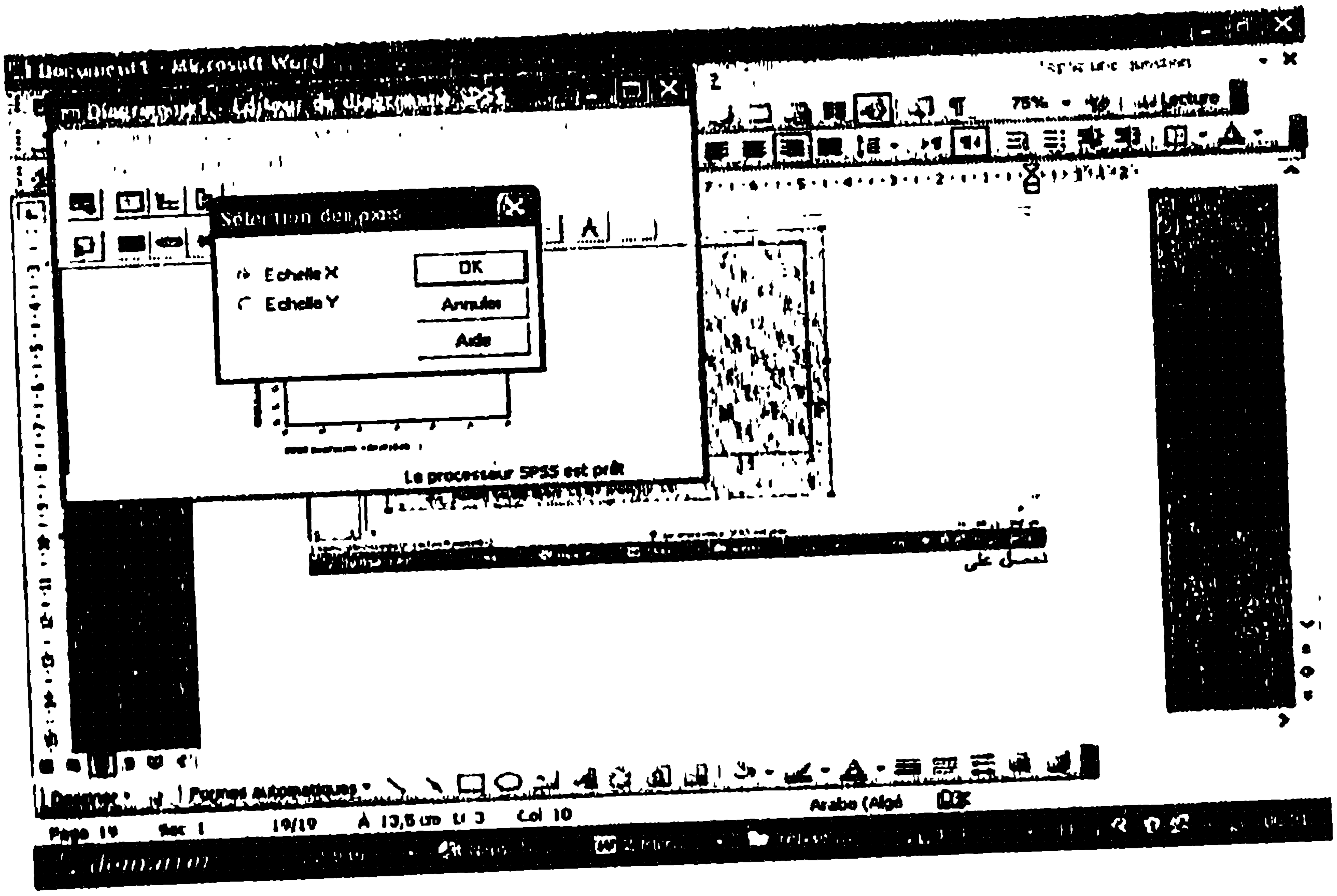
31. نقـوم الآن بتحديـد الإسـقاط بالـضغط مـرتين باسـتعمال المـشيرة La sourie، نحصل على :



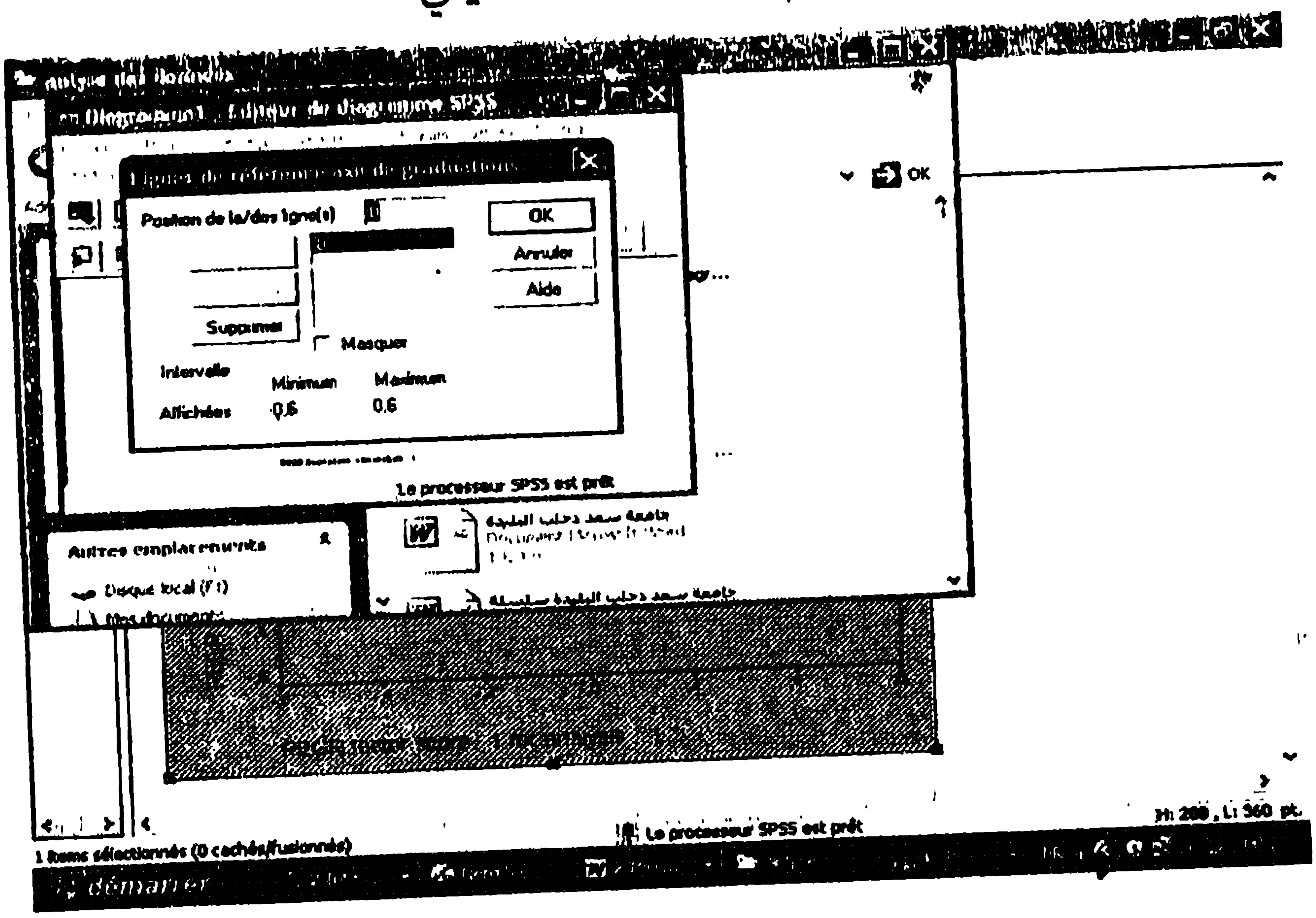
32. نذهب إلى Graphiques، ونقوم بضغط على Ligne de réference



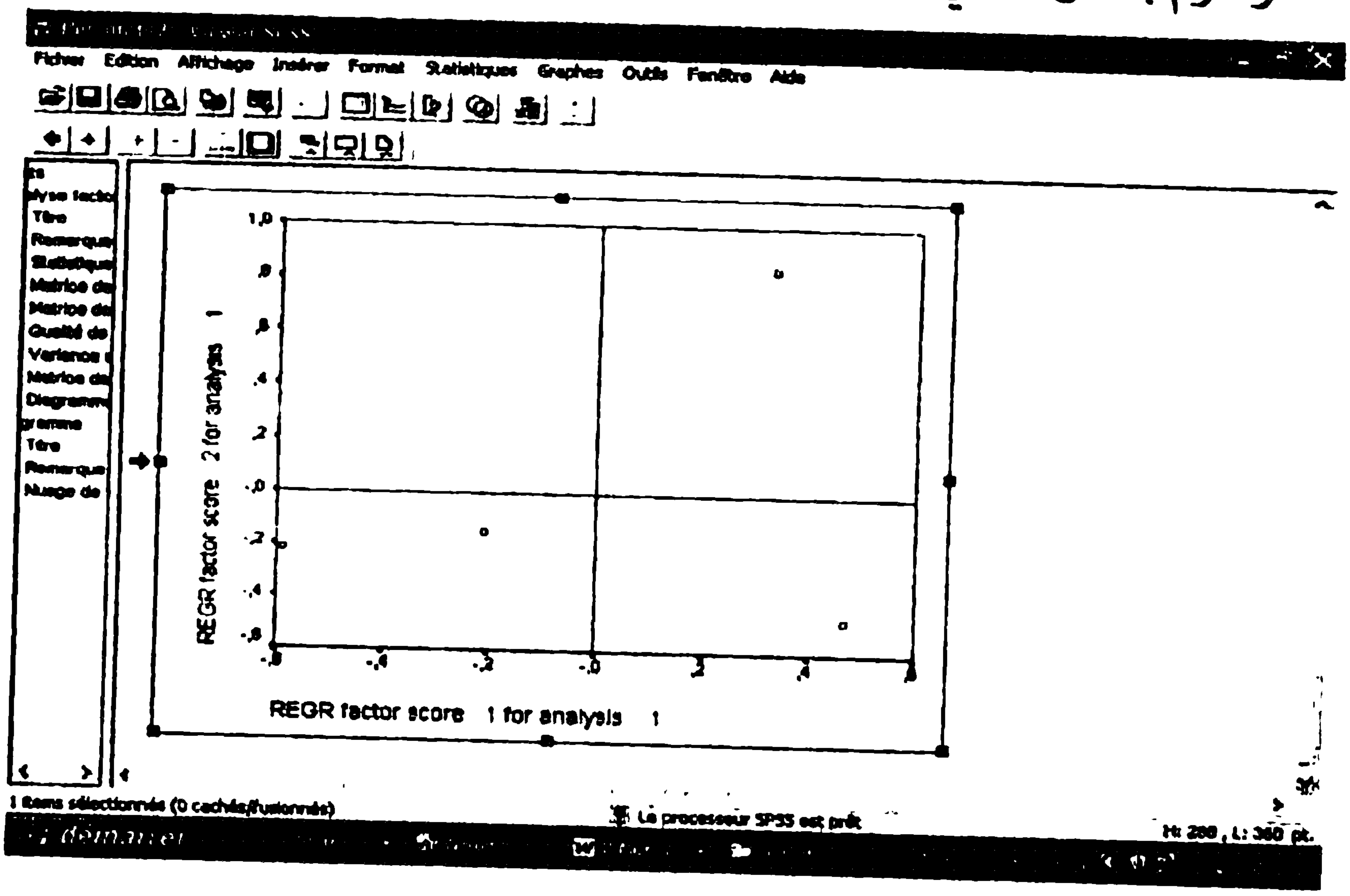
33. نحصل على :



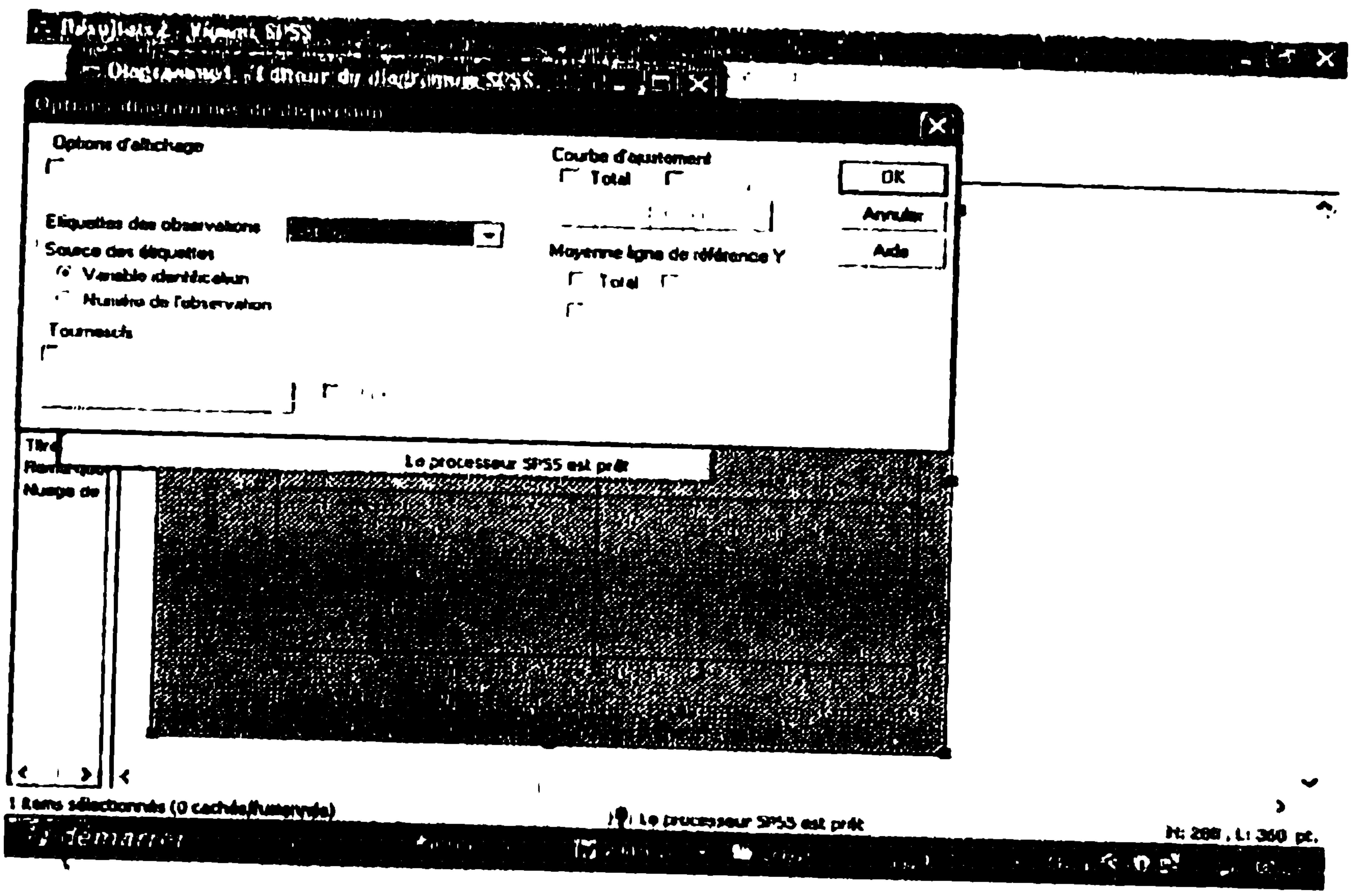
34. بالضغط على Ok، نقوم بتحديد 0 كما يلي :



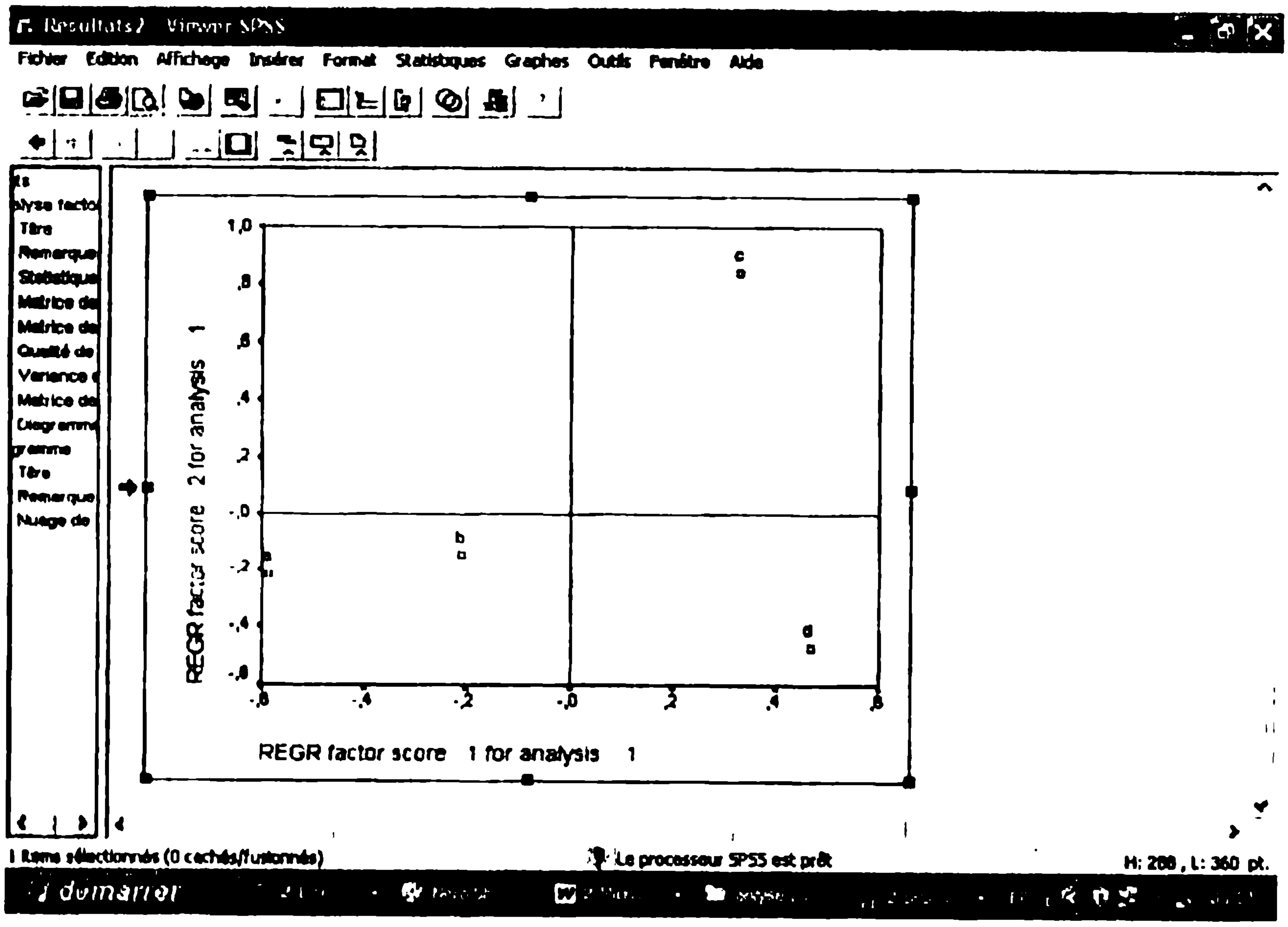
35. ونقوم بنفس الشيء بالنسبة لـ Y نحصل على :



36. نذهب إلى Option، ونقوم بتحديد Activer كما يلي :



37. وبالضغط على OK نحصل على :



ملاحظة : إذا أردنا القيام تحليل المعطيات بالمركبات الرئيسية المعيارية، نتبع نفس الخطوات إلا أننا نختار مصفوفة الارتباطات Matrice des corrélations في المرحلة 19.

الـملحق الثالـث

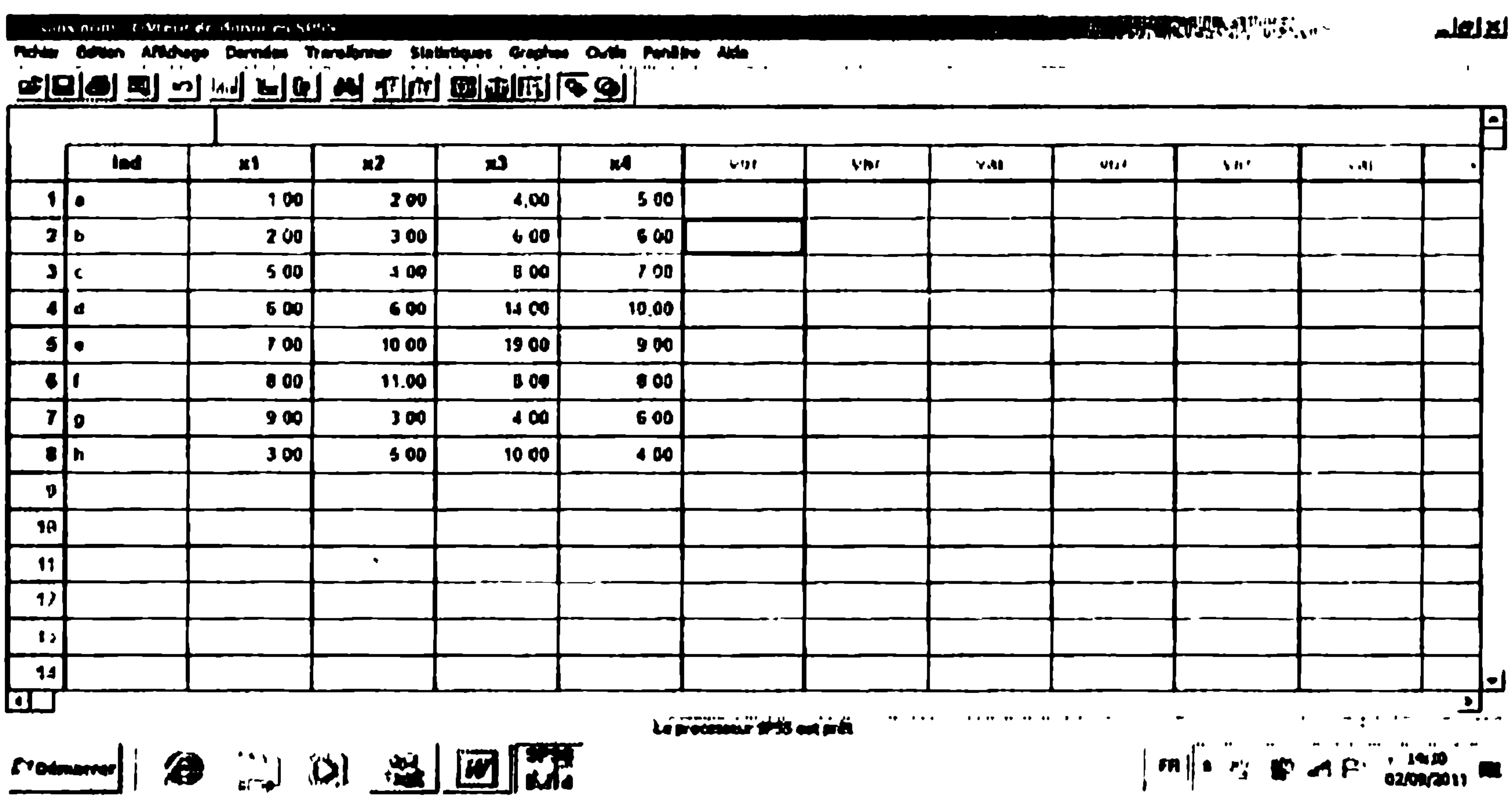
التصنيـف التسلسلـي باستعمـال SPSS

مثـــال :

i	X1	X2	X3	X4
A	1	2	4	5
B	2	3	6	6
C	5	4	8	7
D	6	6	14	10
E	7	10	19	9
F	8	11	8	8
G	9	3	4	6
H	3	5	10	4

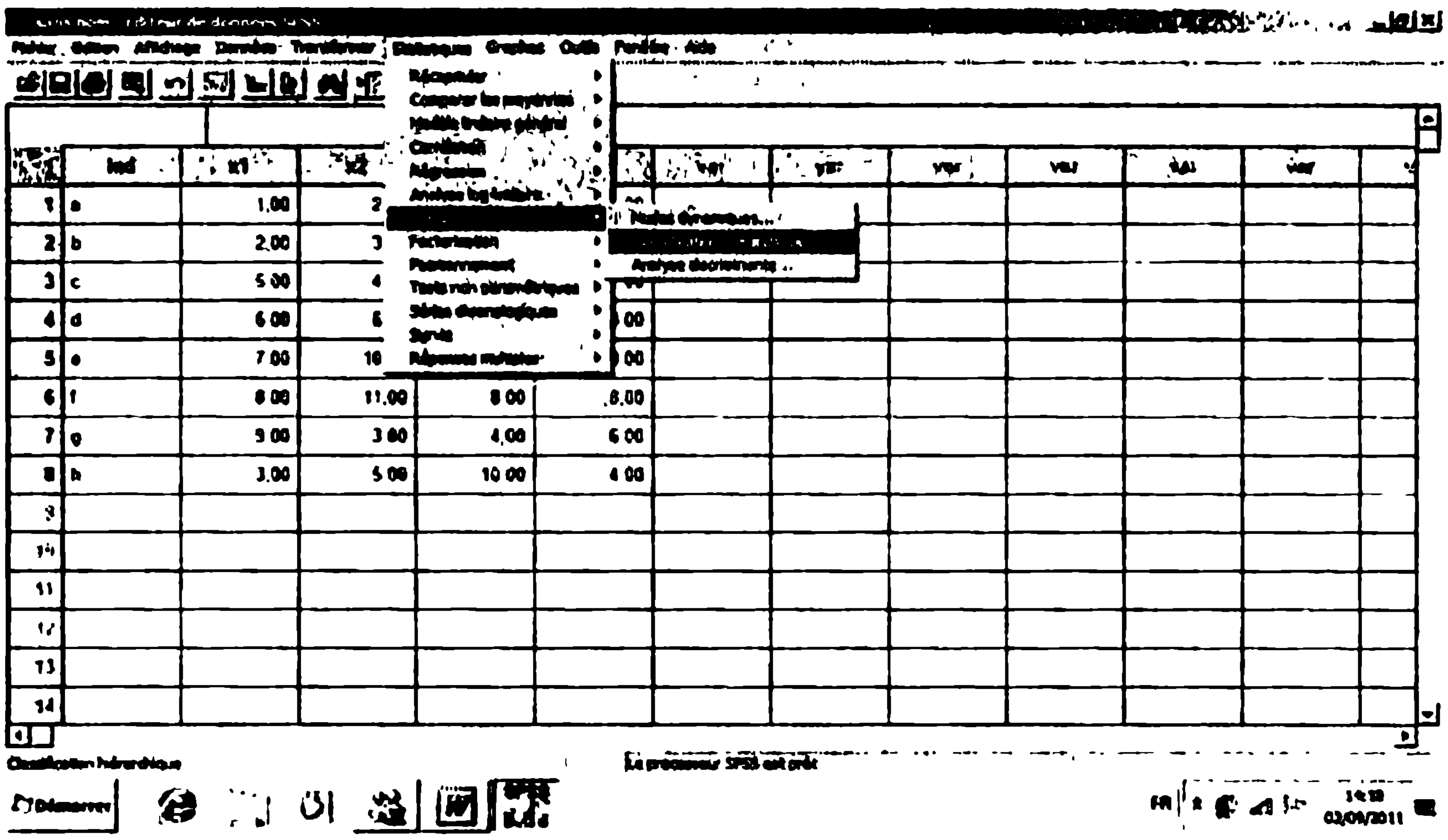
- تصنيف التسلسلي للأفراد باستعمال طريقة WARD.

بعـد تثبيـت برنـامج SPPS 8.0، وادخـال المعطيـات باسـتعمال مراحل (1 إلى 13) نحصل على :

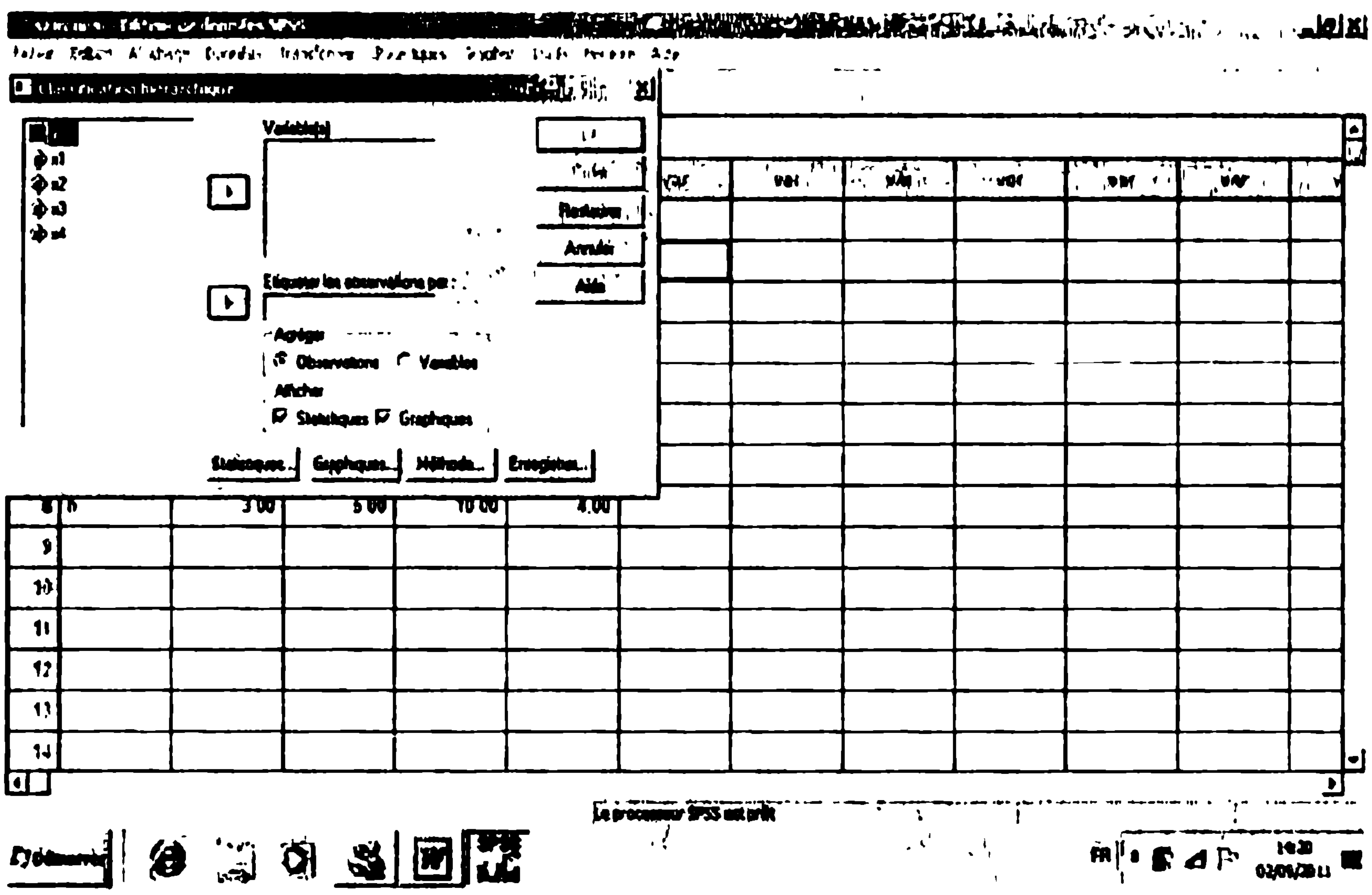


	ind	x1	x2	x3	x4
1	a	1.00	2.00	4.00	5.00
2	b	2.00	3.00	6.00	6.00
3	c	5.00	4.00	8.00	7.00
4	d	6.00	6.00	14.00	10.00
5	e	7.00	10.00	19.00	9.00
6	f	8.00	11.00	8.00	8.00
7	g	9.00	3.00	4.00	6.00
8	h	3.00	5.00	10.00	4.00

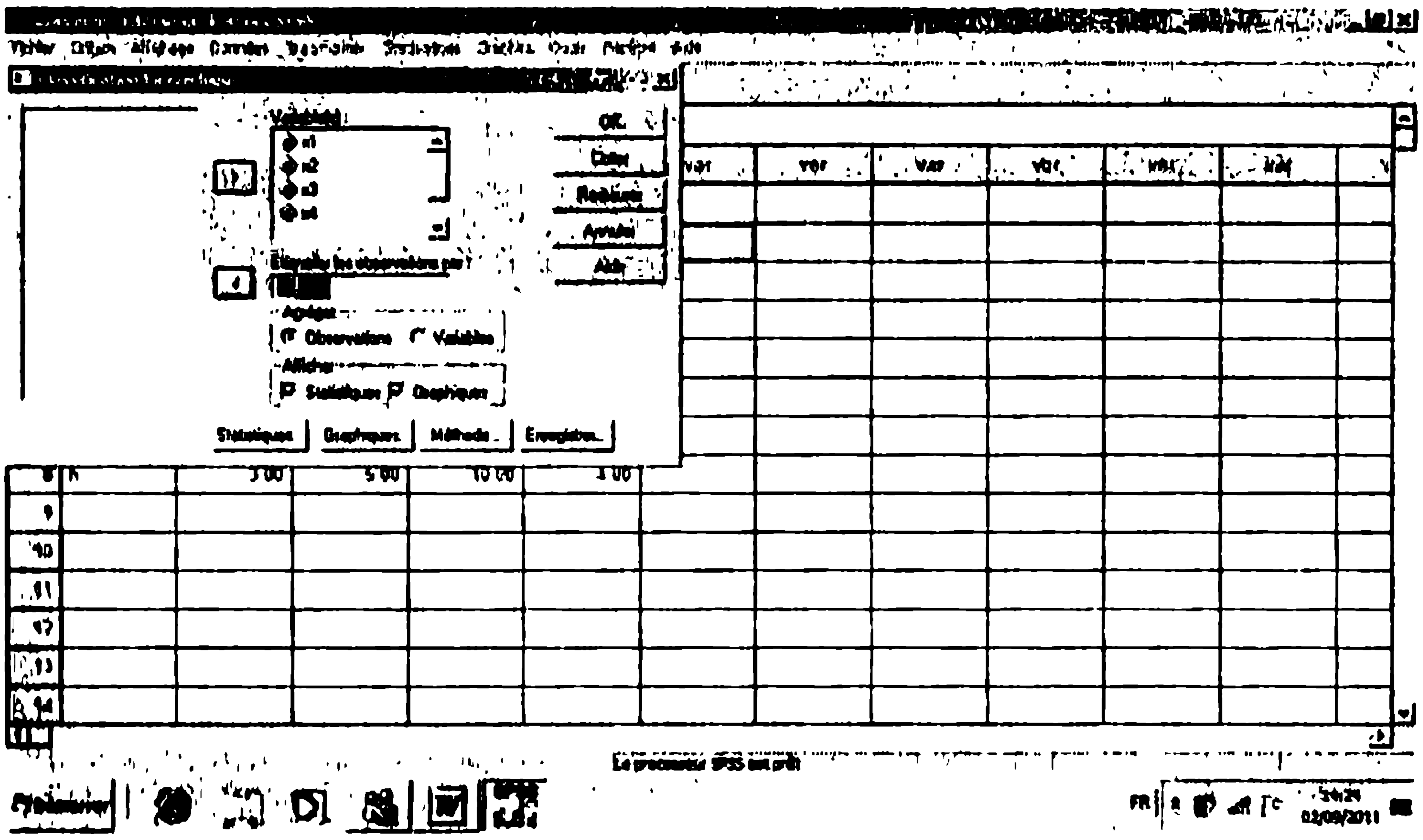
1. نـــذهب الآن إلى Statistiques ، ثـــم Classification، ثـــم علـــى Classification hiérarchique كما يلـي :



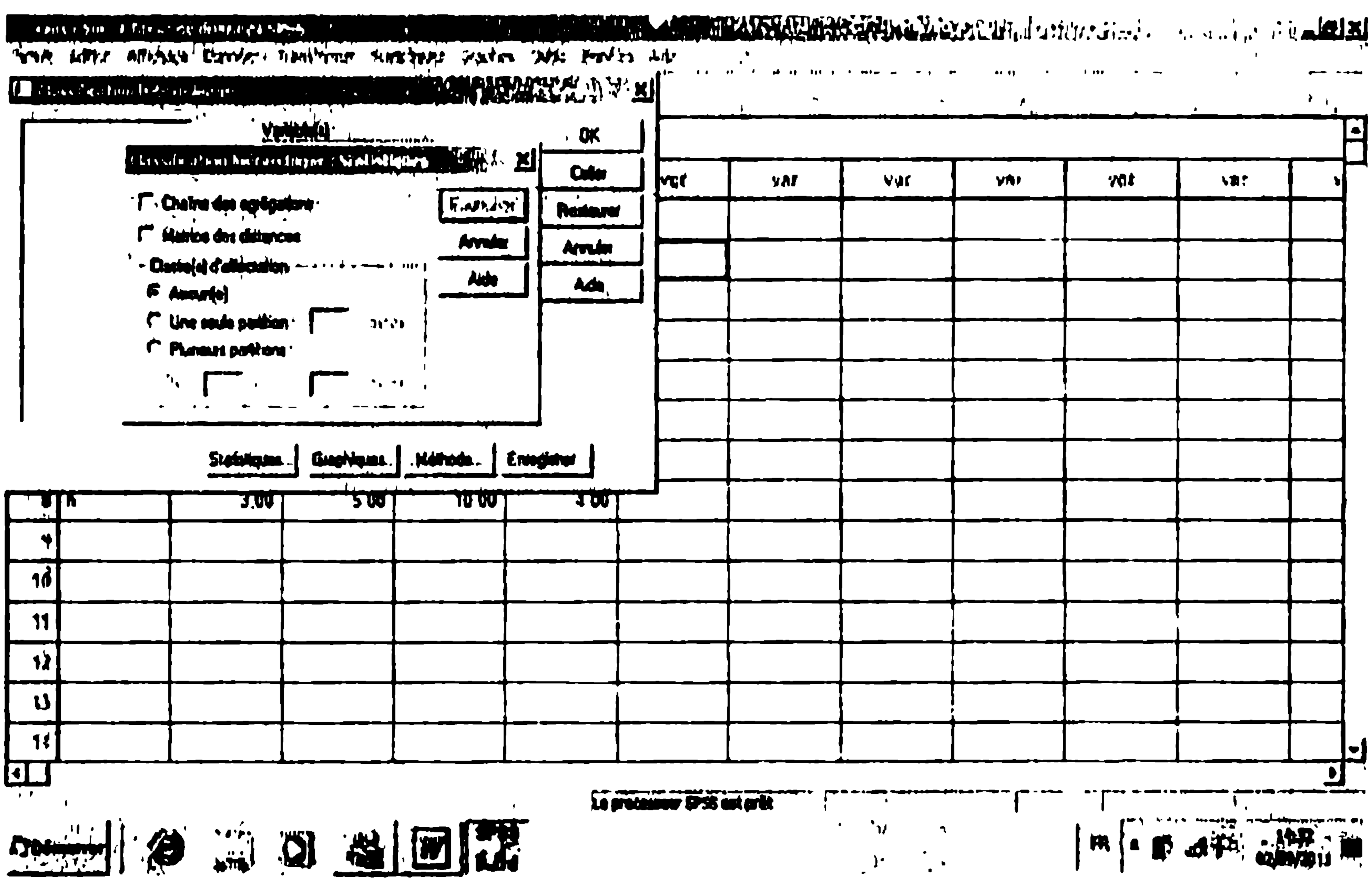
2. بالضغط على Classification hiérarchique نحصل على :



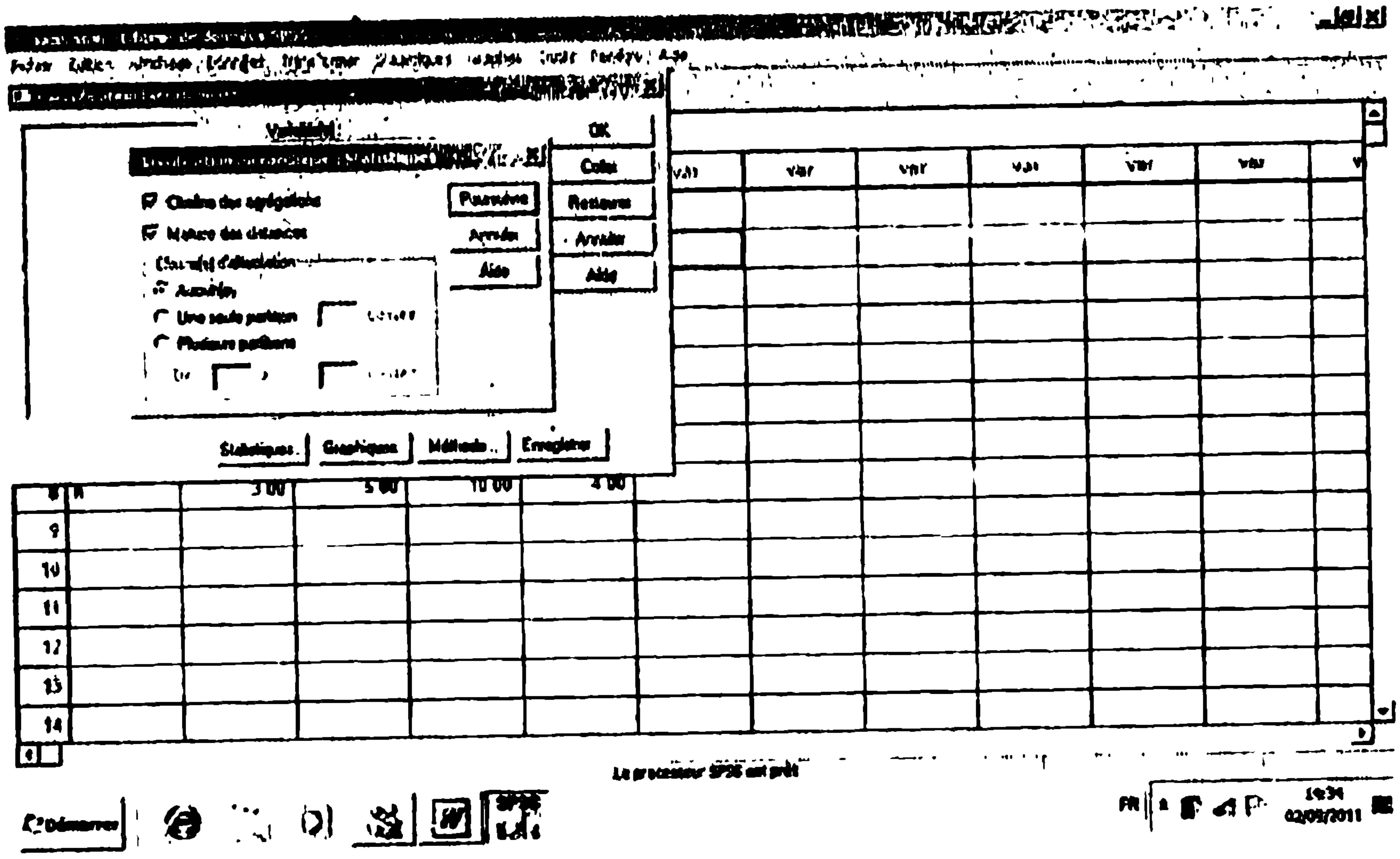
3. نقـوم بإدخـال المـتغيرات X1 ،X2 ،X3 ،X4 في Variables ؛ وind في
: Etiqueter les observations par



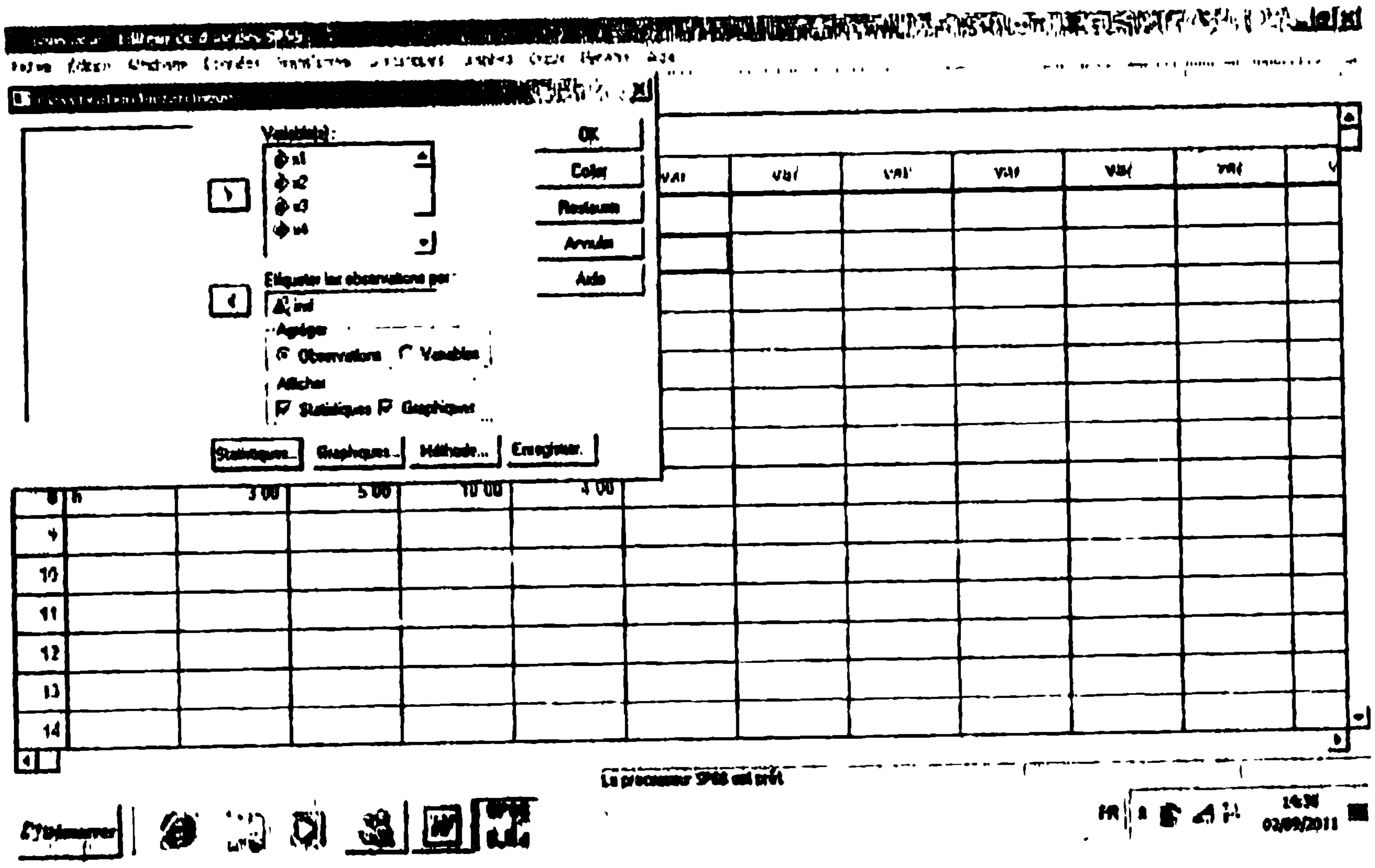
4. نضغط على ... Statistiques نحصل على :



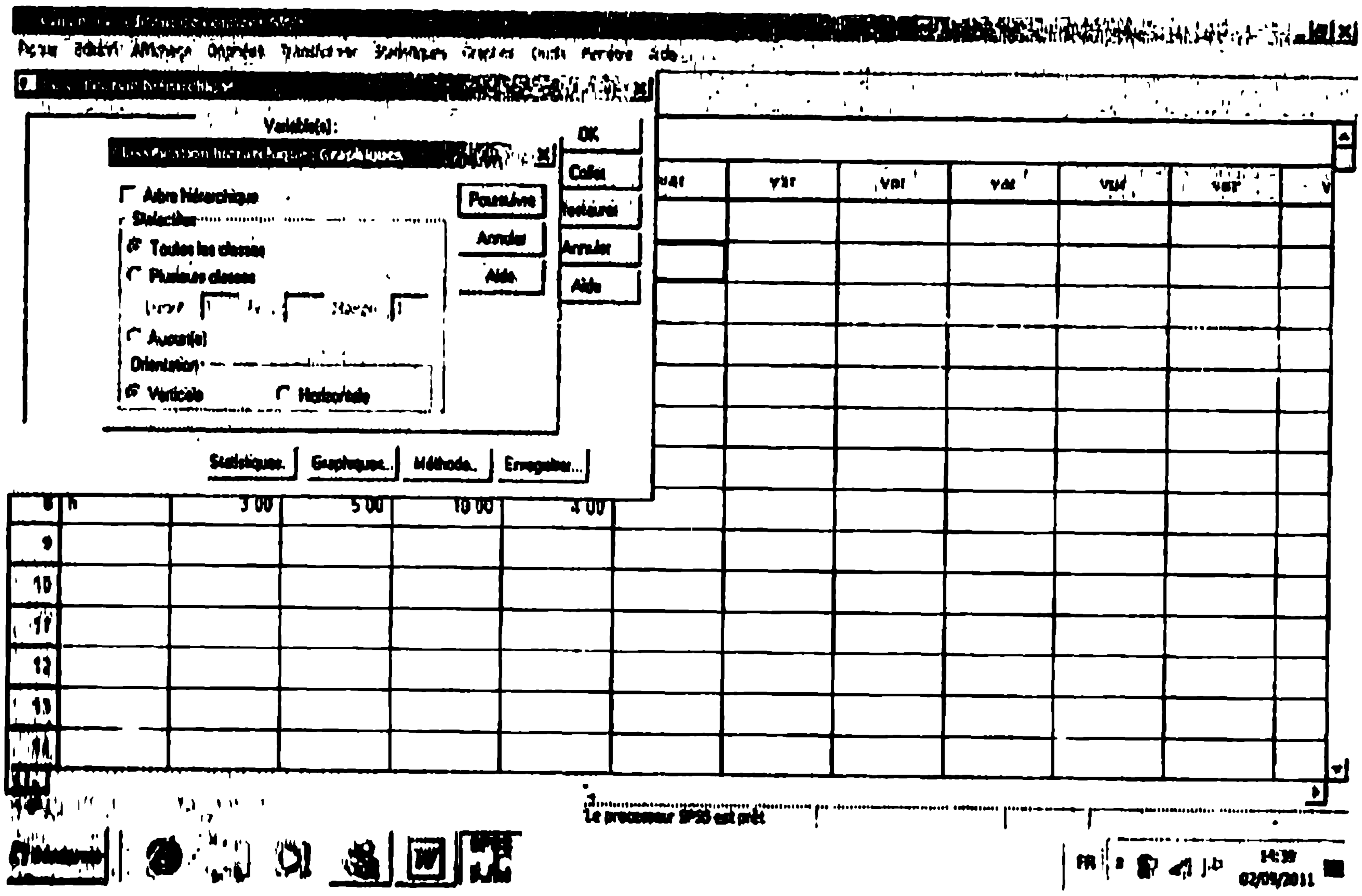
نقوم بتحديد Chaine des aggregations و Matrice des distances كما يلي :



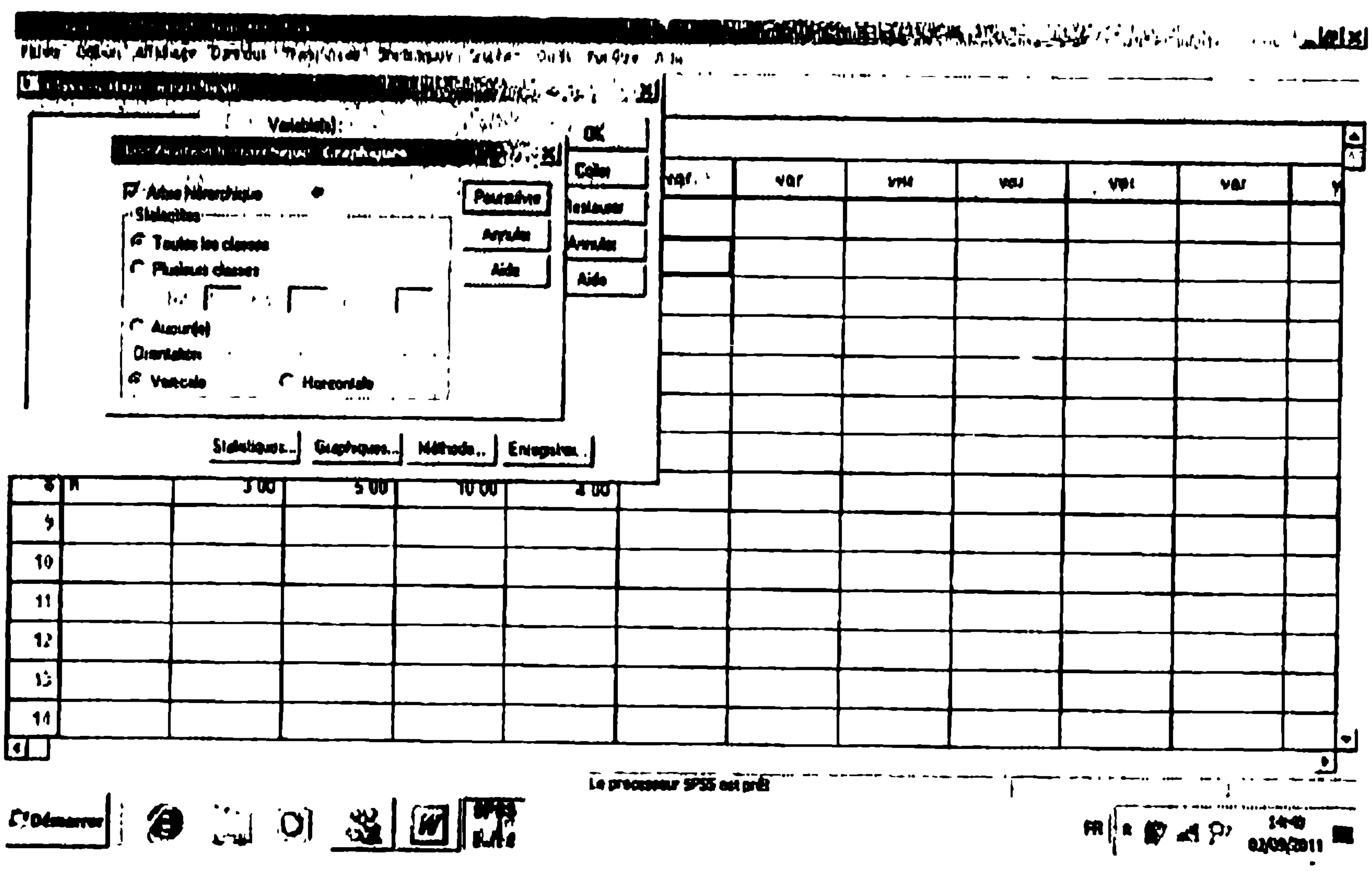
ثم نقوم بالضغط على Poursuivre نحصل على :



5. نضغط على ... Graphiques نحصل على :

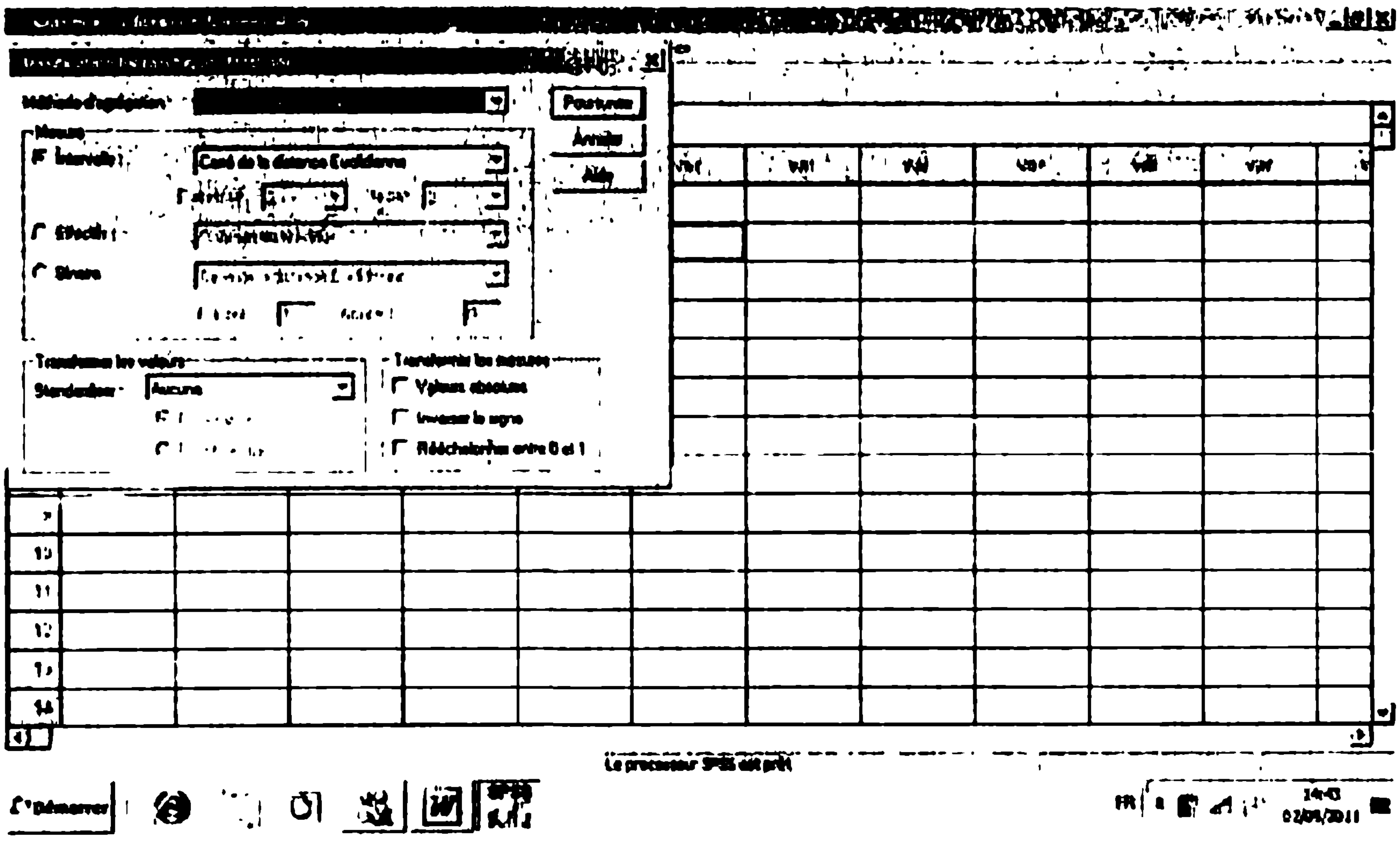


ثم نقوم بتحديد Arbre hiéarchique نحصل على :

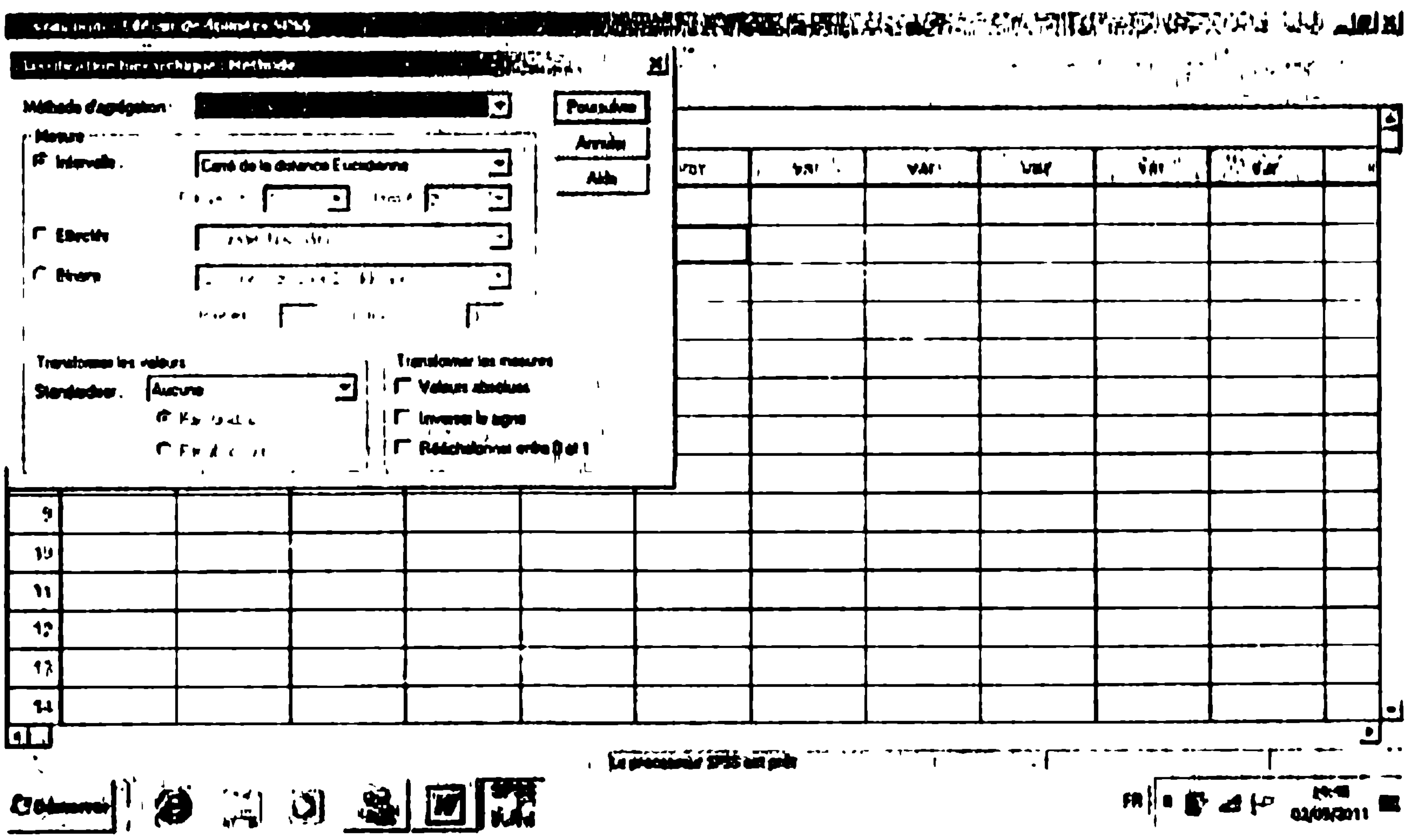


ثم نضغط على Poursuivre

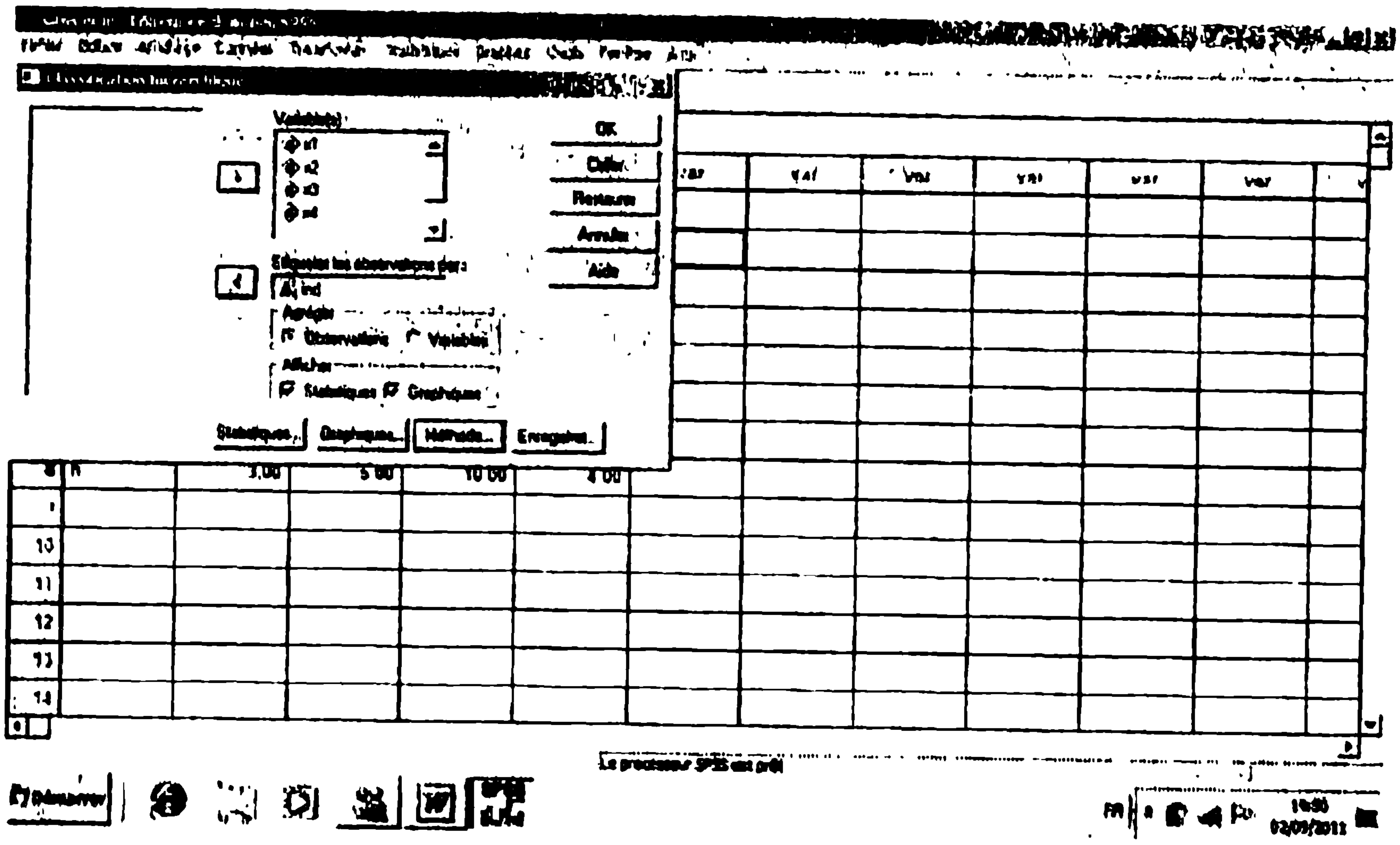
5. نذهب إلى ...Methode، نحصل على :



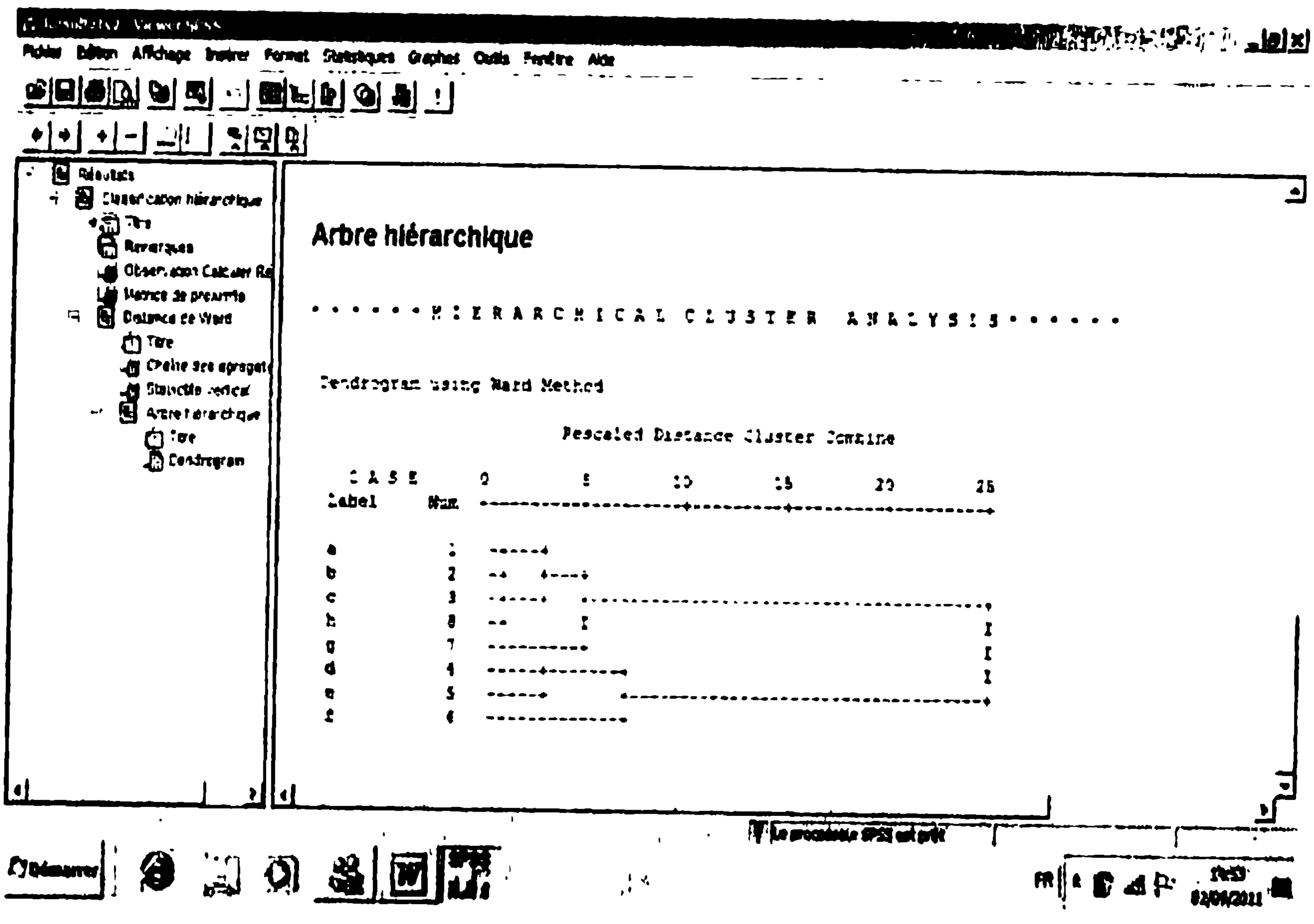
في Méthode d'agrégation نختار طريقة التجميع Méthode Ward في Intervalle نختار Carré de la distance Euclidienne نحصل على :



نضغط على Poursuivre ، فنحصل على :



6. نضغط على OK نحصل علــى :



ملاحظة : إذا أردنا تصنيف المتغيرات نحدد في المرحلة 2، المتغيرات ؛ أما إذا أردنا تغيير طريقة التجميع نذهب إلى المرحلة 4 ونختار طريقة التجميع.

الـــملحق الرابـــع

جـدول كـاي تـربيـــع X2

جدول χ^2 ذات احتمال P و درجة حرية V

P	0,99	0,98	0,95	0,50	0,30	0,20	0,10	0,05	0,025	0,01
v										
1	0,00	0,00	0,00	0,45	1,07	1,64	2,71	3,84	5,02	6,63
2	0,02	0,04	0,10	1,39	2,41	3,22	4,61	5,99	7,38	9,21
3	0,11	0.18	0,35	2,37	3,66	4,64	6,25	7,81	9,35	11,34
4	0,30	0,43	0,71	3,36	4,88	5,99	7,78	9,49	11,14	13,28
5	0,55	0,75	1.15	4,35	6,06	7,29	9,24	11,07	12,83	15,09
6	0,87	1.13	1,64	5,35	7,23	8,56	10,64	12,59	14,45	16,81
7	1,24	1,56	2.17	6,35	8,38	9,80	12,02	14,07	16,01	18,48
8	1,65	2,03	2.73	7,34	9,52	11,03	13.36	15,51	17,53	20,09
9	2,09	2.53	3.33	8,34	10,66	12,24	14,68	16,92	19,02	21,67
10	2,56	3,06	3,94	9,34	11.78	13.44	15.99	18,31	20,48	23,21
11	3,05	3,61	4,57	10,34	12,90	14,63	17,28	19,68	21,92	24,72
12	3,57	4,18	5,23	11,34	14,01	15,81	18,55	21,03	23,34	26,22
13	4.11	4,77	5,89	12,34	15,12	16.98	19,81	22,36	24,74	27,69
14	4,66	5,37	6,57	13,34	16,22	18,15	21,06	23,68	26,12	29,14
15	5,23	5,98	7,26	14,34	17,32	19,31	22,31	25,00	27,49	30,58
16	5,81	6,61	7,96	15,34	18,42	20,47	23.54	26,30	28,85	32,00
17	6,41	7,26	8,67	16,34	19,51	21,61	24,77	27,59	30,19	33,41
18	7,01	7,91	9,39	17,34	20,60	22,76	25,99	28,87	31,53	34,81
19	7,63	8,57	10.12	18,34	21,69	23,90	27,20	30,14	32,85	36,19
20	8,26	9,24	10,85	19,34	22,77	25,04	28,41	31,41	34,17	37,57
21	8.90	9,91	11,59	20,34	23,86	26,17	29.62	32,67	35,48	38,93
22	9,54	10,60	12,34	21,34	24,94	27,30	30,81	33,92	36,78	40,29
23	10,20	11,29	13,09	22,34	26,02	28,43	32,01	35,17	38,08	41,64

تابع لجدول χ^2 ذات احتمال P و درجة حرية V

P	0,99	0,98	0,95	0,50	0,30	0,20	0,10	0,05	0,025	0,01
v										
24	10,86	11,99	13,85	23,34	27,10	29,55	33,20	36,42	39,36	42,98
25	11,52	12,70	14,61	24,34	28,17	30,68	34,38	37,65	40,65	44,31
26	12,20	13,41	15,38	25,34	29,25	31,79	35,56	38,89	41,92	45,64
27	12,88	14,13	16,15	26,34	30,32	32,91	36,74	40,11	43,19	46,96
28	13,56	14,85	16,93	27,34	31,39	34,03	37,92	41,34	44,46	48,28
29	14,26	15,57	17,71	28,34	32,46	35,14	39,09	42,56	45,72	49,59
30	14,95	16,31	18,49	29,34	33,53	36,25	40,26	43,77	46,98	50,89
31	15,66	17,04	19,28	30,34	34,60	37,36	41,42	44,99	48,23	52,19
32	16,36	17,78	20,07	31,34	35,66	38,47	42,58	46,19	49,48	53,49
33	17,07	18,53	20,87	32,34	36,73	39,57	43,75	47,40	50,73	54,78
34	17,79	19,28	21,66	33,34	37,80	40,68	44,90	48,60	51,97	56,06
35	18,51	20,03	22,47	34,34	38,86	41,78	46,06	49,80	53,20	57,34
36	19,23	20,78	23,27	35,34	39,92	42,88	47,21	51,00	54,44	58,62
37	19,96	21,54	24,07	36,34	40,98	43,98	48,36	52,19	55,67	59,89
38	20,69	22,30	24,88	37,34	42,05	45,08	49,51	53,38	56,90	61,16
39	21,43	23,07	25,70	38,34	43,11	46,17	50,66	54,57	58,12	62,43
40	22,16	23,84	26,51	39,34	44,16	47,27	51,81	55,76	59,34	63,69
41	22,91	24,61	27,33	40,34	45,22	48,36	52,95	56,94	60,56	64,95
42	23,65	25,38	28,14	41,34	46,28	49,46	54,09	58,12	61,78	66,21
43	24,40	26,16	28,96	42,34	47,34	50,55	55,23	59,30	62,99	67,46
44	25,15	26,94	29,79	43,34	48,40	51,64	56,37	60,48	64,20	68,71
45	25,90	27,72	30,61	44,34	49,45	52,73	57,51	61,66	65,41	69,96
46	26,66	28,50	31,44	45,34	50,51	53,82	58,64	62,83	66,62	71,20
47	27,42	29,29	32,27	46,34	51,56	54,91	59,77	64,00	67,82	72,44
48	28,18	30,08	33,10	47,34	52,62	55,99	60,91	65,17	69,02	73,68
49	28,94	30,87	33,93	48,33	53,67	57,08	62,04	66,34	70,22	74,92
50	29,71	31,66	34,76	49,33	54,72	58,16	63,17	67,50	71,42	76,15
51	30,48	32,46	35,60	50,33	55,78	59,25	64,30	68,67	72,62	77,39
52	31,25	33,26	36,44	51,33	56,83	60,33	65,42	69,83	73,81	78,62

تابع لجدول χ^2 ذات احتمال P و درجة حرية V

P v	0,99	0,98	0,95	0,50	0,30	0,20	0,10	0,05	0,025	0,01
53	32,02	34,06	37,28	52,33	57,88	61,41	66,55	70,99	75,00	79,84
54	32,79	34,86	38,12	53,33	58,93	62,50	67,67	72,15	76,19	81,07
55	33,57	35,66	38,96	54,33	59,98	63,58	68,80	73,31	77,38	82,29
56	34,35	36,46	39,80	55,33	61,03	64,66	69,92	74,47	78,57	83,51
57	35,13	37,27	40,65	56,33	62,08	65,74	71,04	75,62	79,75	84,73
58	35,91	38,08	41,49	57,33	63,13	66,82	72,16	76,78	80,94	85,95
59	36,70	38,89	42,34	58,33	64,18	67,89	73,28	77,93	82,12	87,17
60	37,48	39,70	43,19	59,33	65,23	68,97	74,40	79,08	83,30	88,38
61	38,27	40,51	44,04	60,33	66,27	70,05	75,51	80,23	84,48	89,59
62	39,06	41,33	44,89	61,33	67,32	71,13	76,63	81,38	85,65	90,80
63	39,86	42,14	45,74	62,33	68,37	72,20	77,75	82,53	86,83	92,01
64	40,65	42,96	46,59	63,33	69,42	73,28	78,86	83,68	88,00	93,22
65	41,44	43,78	47,45	64,33	70,46	74,35	79,97	84,82	89,18	94,42
66	42,24	44,60	48,31	65,33	71,51	75,42	81,09	85,96	90,35	95,63
67	43,04	45,42	49,16	66,33	72,55	76,50	82,20	87,11	91,52	96,83
68	43,84	46,24	50,02	67,33	73,60	77,57	83,31	88,25	92,69	98,03
69	44,64	47,07	50,88	68,33	74,64	78,64	84,42	89,39	93,86	99,23
70	45,44	47,89	51,74	69,33	75,69	79,71	85,53	90,53	95,02	100,43
71	46,25	48,72	52,60	70,33	76,73	80,79	86,64	91,67	96,19	101,62
72	47,05	49,55	53,46	71,33	77,78	81,86	87,74	92,81	97,35	102,82
73	47,86	50,38	54,33	72,33	78,82	82,93	88,85	93,95	98,52	104,01
74	48,67	51,21	55,19	73,33	79,86	84,00	89,96	95,08	99,68	105,20
75	49,48	52,04	56,05	74,33	80,91	85,07	91,06	96,22	100,84	106,39
76	50,29	52,87	56,92	75,33	81,95	86,13	92,17	97,35	102,00	107,58
77	51,10	53,71	57,79	76,33	82,99	87,20	93,27	98,48	103,16	108,77
78	51,91	54,54	58,65	77,33	84,04	88,27	94,37	99,62	104,32	109,96
79	52,72	55,38	59,52	78,33	85,08	89,34	95,48	100,75	105,47	111,14
80	53,54	56,21	60,39	79,33	86,12	90,41	96,58	101,88	106,63	112,33

تابع لجدول χ^2 ذات احتمال P و درجة حرية V

P v	0,99	0,98	0,95	0,50	0,30	0,20	0,10	0,05	0,025	0,01
81	54,36	57,05	61,26	80,33	87,16	91,47	97,68	103,01	107,78	113,51
82	55,17	57,89	62,13	81,33	88,20	92,54	98,78	104,14	108,94	114,69
83	55,99	58,73	63,00	82,33	89,24	93,60	99,88	105,27	110,09	115,88
84	56,81	59,57	63,88	83,33	90,28	94,67	100,98	106,39	111,24	117,06
85	57,63	60,41	64,75	84,33	91,32	95,73	102,08	107,52	112,39	118,24
86	58,46	61,25	65,62	85,33	92,37	96,80	103,18	108,65	113,54	119,41
87	59,28	62,10	66,50	86,33	93,41	97,86	104,28	109,77	114,69	120,59
88	60,10	62,94	67,37	87,33	94,44	98,93	105,37	110,90	115,84	121,77
89	60,93	63,79	68,25	88,33	95,48	99,99	106,47	112,02	116,99	122,94
90	61,75	64,63	69,13	89,33	96,52	101,05	107,57	113,15	118,14	124,12
91	62,58	65,48	70,00	90,33	97,56	102,12	108,66	114,27	119,28	125,29
92	63,41	66,33	70,88	91,33	98,60	103,18	109,76	115,39	120,43	126,46
93	64,24	67,18	71,76	92,33	99,64	104,24	110,85	116,51	121,57	127,63
94	65,07	68,03	72,64	93,33	100,68	105,30	111,94	117,63	122,72	128,80
95	65,90	68,88	73,52	94,33	101,72	106,36	113,04	118,75	123,86	129,97
96	66,73	69,73	74,40	95,33	102,76	107,43	114,13	119,87	125,00	131,14
97	67,56	70,58	75,28	96,33	103,79	108,49	115,22	120,99	126,14	132,31
98	68,40	71,43	76,16	97,33	104,83	109,55	116,32	122,11	127,28	133,48
99	69,23	72,29	77,05	98,33	105,87	110,61	117,41	123,23	128,42	134,64
100	70,06	73,14	77,93	99,33	106,91	111,67	118,50	124,34	129,56	135,81

قائمة المراجع

- Cordan, A. D. (1999), Classification, Chapman & Hall, London.

- De Lagarde, J. (1998), Initiation à l'analyse des données, Dunod, 3é, Paris.

- Diday, E. & Lemaire. J. & Pouget, J. & Testu, F. (1982), Éléments d'analyse de données. Bordas, Paris.

- Escofier, B, Pagès. J. (1998), Analyses factorielles simple et multiples. 3éditon, Dunod, Paris

- Everitt, B. S. (1977), Graphical Techniques for Multivariate Data, Published by Heinemann Educational Books Ltd, London.

- Everitt. B. S. (1993), Cluster Analysis, Published by Heinemann Educational Books Ltd,London, and Copublished in the America by Halsted Press, New York.

- Flury, B. & Riedwyl, H. (1993), Multivariate Statistics a practical approach, Chapman & Hall, London.

- Ghidaglia, J. M. & Guyon, X. (2000), Mathématiques & Applications, Springer, Paris.

- Jambu, M. (1999), Méthodes de base de l'analyse des données, éditions Eyrolles, Paris

-Lebart, L. Morineau, A. Piron, M. (2000), Statistique exploratoire multidimensionnelle, 3 édition, Dunod, Paris.

- Nakache, J-P. Chevalier, A. Morice, V. (1981), Exercices commentés de mathématiques pour l'analyse statistique des données, Bordas, Paris.

- Tenenhaus, M. (2007), Statistique (Méthodes pour décrire, expliquer et prévoir), Dunod, Paris.

- Zupan, J. (1982), Clustering of Large Data Sets, British Library Cataloguing in Publication Data.

طبع بمطبعة دار هومه – الجزائر 2012
34، حي لابرويار – بوزريعة – الجزائر
الهاتف: 021.94.41.19/ 021.94.19.36
الفاكس: 021.94.17.75/ 021.79.91.84
www.editionshouma.com
email : Info@editionshouma.com

ISBN: 978-9961-65-619-8 :ردمك

دار هومه
للطباعة والنشر والتوزيع
34 حي لابروفيار - بوزريعة - الجزائر
الفاكس: 75 17 94 021 الهاتف: 36 19 94 021
021 79 91 84 021 94 41 19